

91759

From: Mertz, Prema
Sent: Thursday, April 17, 2003 1:35 PM
To: STIC-Biotech/ChemLib
Subject: 09/874,056

Please search SEQ ID NO:4 with protein databases.

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

10019

RECEIVED
APR 17 2003
(STIC)

Point of Contact
Searcher: P. Sheppard
Phone: Telephone number: (703) 308-4499
Location: _____
Date Picked Up: _____
Date Completed: 4/22/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:12:05 ; Search time 28 seconds
(without alignments)

84.434 Million cell updates/sec

Title: US-09-874-056-4

Perfect score: 352

Sequence: 1 IGCDDHTSPVGTGCTCPSLK.....QHCCPAGYTCNVKARTCEKQ 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	347	98.6	588	1 GRN_RAT	P23785 r granulins
2	347	98.6	589	1 GRN_MOUSE	P28798 mus musculus
3	331	94.0	593	1 GRN_HUMAN	P28799 h granulins
4	307	87.2	591	1 GRN_CAVPO	P28797 cavia porcea
5	158	44.9	346	1 CYSL_LYCES	P20721 lycopersico
6	150	42.6	471	1 ORYB_ORYSA	P25777 oryza sativ
7	145	41.2	57	1 GRN1_CYPCA	P81013 cyprinus ca
8	144	40.9	57	1 GRN3_CYPCA	P81015 cyprinus ca
9	140	39.8	57	1 GRN2_CYPCA	P81014 cyprinus ca
10	136	38.6	458	1 ORYA_ORYSA	P25776 oryza sativ
11	131.5	37.4	54	1 PMD1_LOCMI	P80059 locusta mig
12	130	36.9	462	1 RD21_ARATH	P43297 arabidopsis
13	128	36.4	46	1 ENAL_HORSE	P80930 equus caball
14	93.5	26.6	116	1 MCS_HUMAN	P49901 homo sapien
15	84.5	24.0	676	1 PRTS_HUMAN	P07225 homo sapien
16	83.5	23.7	714	1 DLL1_RAT	P97677 rattus norv
17	82.5	23.4	649	1 PRTS_MACMU	Q28520 macaca mula
18	82.5	23.4	685	1 DLL4_HUMAN	Q9nr61 homo sapien
19	82.5	23.4	722	1 DLL1_MOUSE	Q61483 mus musculus
20	81	23.0	61	1 MTA_ONCMY	P09861 oncorhynch
21	80.5	22.9	686	1 DLL4_MOUSE	Q9j171 mus musculus
22	80.5	22.9	723	1 DLL1_HUMAN	O00548 homo sapien
23	79	22.4	4660	1 LRP2_RAT	P98158 rattus norv
24	78	22.2	61	1 MTLA_BOVIN	P04356 bos taurus
25	77.5	22.0	1246	1 YWV2_CAEEL	P34504 caenorhabdi
26	77	21.9	61	1 MTLA_PIG	P49068 sus scrofa
27	77	21.9	638	1 MT_CHICK	P09576 gallus gall
28	76.5	21.7	1328	1 AGRI_DISOM	Q90404 discopoge o
29	75.5	21.4	74	1 MTL_CAEEL	P17511 caenorhabdi
30	75	21.3	61	1 MTLF_PIG	P79378 sus scrofa
31	74.5	21.2	60	1 MT_GADMO	P51902 gadus morhu
32	74.5	21.2	705	1 FBL1_MOUSE	Q08879 mus musculus
33	74	21.0	61	1 MTL_E_PIG	P79431 sus scrofa

RESULT 1

GRN_RAT ID	GRN_RAT	STANDARD	PRT	588 AA
AC	P23785;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Granulins precursor (Acrogranin) [Contains: Granulin 1 (Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B) (Epithelin 2); Granulin 4 (Granulin A) (Epithelin 1); Granulin 5 (Granulin C); Granulin 6 (Granulin D); Granulin 7 (Granulin E)].			
GN	GRN.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=94062640; PubMed=8243292;			
RA	Bhandari V., Giald A., Bateman A.;			
RT	"The complementary deoxyribonucleic acid sequence, tissue distribution, and cellular localization of the rat granulin precursor.";			
RT	Endocrinology 133:2682-2689(1993).			
RL	[2]			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 204-259 AND 278-334.			
RP	TISSUE=Kidney;			
RC	MEDLINE=92317004; PubMed=1618805;			
RX	Plowman G.D., Green J.M., Neubauer M.G., Buckley S.D., McDonald V.L.,			
RA	Todaro G.J., Shoyab M.;			
RT	"The epithelin precursor encodes two proteins with opposing activities on epithelial cell growth.";			
RT	J. Biol. Chem. 267:13073-13078(1992).			
RN	[3]			
RP	SEQUENCE OF 204-225 AND 279-299.			
RX	MEDLINE=91045907; PubMed=2236009;			
RA	Shoyab M., McDonald V.L., Byles C., Todaro G.J., Plowman G.D.;			
RT	"Epithelins 1 and 2: isolation and characterization of two cysteine-rich growth-modulating proteins.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 87:7912-7916(1990).			
RL	[4]			
RN	SEQUENCE OF 278-328.			
RP	TISSUE=Bone marrow;			
RC	MEDLINE=91097544; PubMed=2268320;			
RX	Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;			
RA	"Granulins, a novel class of peptide from leukocytes.";			
RL	Biochem. Biophys. Res. Commun. 173:1161-1168(1990).			
CC	-!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.			
CC	-!- TISSUE SPECIFICITY: UBIQUITOUS; MOST ABUNDANT IN THE SPLEEN AND SEVERAL TISSUES OF ENDOCRINE SIGNIFICANCE.			
CC	-!- PTM: GRANULINS ARE DISULFIDE BRIDGED.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -			

P58280 bos taurus
P02804 cricetus
P15265 mus musculus
P16144 homo sapien
O13258 chionodraco
O73914 notothenia
O93609 pagothenia
O13259 chionodraco
O92145 pagothenia
O93450 parachaenic
P52727 sparus aura
P52721 thermarces

ALIGNMENTS

34 74 21.0 61 1 MTL_BOVIN
35 74 21.0 61 1 MTL_CRIGR
36 74 21.0 197 1 MCS_MOUSE
37 74 21.0 1822 1 ITB4_HUMAN
38 73.5 20.9 60 1 MTA_CHIHA
39 73.5 20.9 60 1 MTA_NOTCO
40 73.5 20.9 60 1 MTA_PAGBE
41 73.5 20.9 60 1 MTB_CHIHA
42 73.5 20.9 60 1 MTB_PAGBE
43 73.5 20.9 60 1 MT_PARCR
44 72.5 20.6 60 1 MTA_SPAAU
45 72.5 20.6 60 1 MTA_THECR

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M97750; AAA16903.1; -;
 DR EMBL; X62322; CAA44198.1; -;
 DR PIR; A36199; A36199.
 DR PIR; B36199; B36199.
 DR PIR; E36698; E36698.
 DR InterPro: IPR000118; Granulin.
 DR Pfam; PF00396; granulin; 7.
 DR SMART; SM00277; GRAN; 7.
 DR PROSITE; PS00799; GRANULINS; 7.
 KW Cytokine; Repeat; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 588 ACGRANIN.
 FT PEPTIDE 58 113 GRANULIN 1.
 FT PEPTIDE 122 178 GRANULIN 2.
 FT PEPTIDE 204 259 GRANULIN 3.
 FT PEPTIDE 278 334 GRANULIN 4.
 FT PEPTIDE 361 413 GRANULIN 5.
 FT PEPTIDE 438 492 GRANULIN 6.
 FT PEPTIDE 512 567 GRANULIN 7.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 201 201 S -> FP (IN REF. 2).
 FT CONFLICT 307 308 TK -> SB (IN REF. 4).
 FT CONFLICT 324 324 Q -> T (IN REF. 4).
 FT CONFLICT 388 388 M -> I (IN REF. 2).
 SQ SEQUENCE 588 AA; 63369 MW; 113D434F7E099B31 CRC64;

Query Match 98.6%; Score 347; DB 1; Length 588;
 Best Local Similarity 100.0%; Pred. No. 4e-28; 0; Indels 0; Gaps 0;
 Matches 56; Conservative 0; Mismatches 0;

QY 1 IGCQHTSCPVGTCPCPSLKGWACCOLPHAVCCEDRQHCPCAGYTCNVKARTCEK 56
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 439 IGCQHTSCPVGTCPCPSLKGWACCOLPHAVCCEDRQHCPCAGYTCNVKARTCEK 494

RESULT 2
 GRN_MOUSE
 ID GRN_MOUSE STANDARD; PRT; 589 AA.
 AC P28798;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2;
 DE Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7].
 GN GRN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9324591; PubMed=8482392;
 RA Baba T., Nemoto H., Watanabe K., Arai Y., Gerton G.L.;
 RT "Exon/intron organization of the gene encoding the mouse
 RT epithelin/granulin precursor (acrogranin).";
 RL FEBS Lett. 322:89-94(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=92317004; PubMed=1618805;
 RA Plowman G.D., Green I.M., Neubauer M.G., Buckley S., McDonald V.L.,
 RA Todaro G.I., Shoyab M.;
 RT "The epithelin precursor encodes two proteins with opposing
 RT activities on epithelial cell growth";
 RL J. Biol. Chem. 267:13073-13078(1992).

CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY
 CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; D16195; BAA03736.1; -;
 DR EMBL; M86736; AAA37191.1; -;
 DR EMBL; X62321; CAA44197.1; -;
 DR MGD; MGI:95832; Grn.
 DR InterPro: IPR000118; Granulin.
 DR Pfam; PF00396; granulin; 7.
 DR SMART; SM00277; GRAN; 7.
 DR PROSITE; PS00799; GRANULINS; 7.
 KW Cytokine; Repeat; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 589 ACGRANIN.
 FT PEPTIDE 758 7113 GRANULIN 1.
 FT PEPTIDE 7122 7178 GRANULIN 2.
 FT PEPTIDE 205 260 GRANULIN 3.
 FT PEPTIDE 280 334 GRANULIN 4.
 FT PEPTIDE 362 7414 GRANULIN 5.
 FT PEPTIDE 440 7493 GRANULIN 6.
 FT PEPTIDE 7517 7568 GRANULIN 7.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 350 350 L -> R (IN REF. 2).
 SQ SEQUENCE 589 AA; 63458 MW; 1DE8229C413CB787 CRC64;

Query Match 98.6%; Score 347; DB 1; Length 589;
 Best Local Similarity 100.0%; Pred. No. 4e-28; 0; Indels 0; Gaps 0;
 Matches 56; Conservative 0; Mismatches 0;

QY 1 IGCQHTSCPVGTCPCPSLKGWACCOLPHAVCCEDRQHCPCAGYTCNVKARTCEK 56
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 440 IGCQHTSCPVGTCPCPSLKGWACCOLPHAVCCEDRQHCPCAGYTCNVKARTCEK 495

RESULT 3
 GRN_HUMAN
 ID GRN_HUMAN STANDARD; PRT; 593 AA.
 AC P28799; P23781; P23782; P23783; P23784; Q9BWE7;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Granulins precursor (Acrogranin) [Contains: Paragranulin; Granulin 1
 DE (Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B);
 DE Granulin 4 (Granulin A); Granulin 5 (Granulin C); Granulin 6 (Granulin
 DE D); Granulin 7 (Granulin E)].
 GN GRN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Bone marrow;
 RX MEDLINE=92179253; PubMed=1542665;
 RA Bhandari V., Palfree R.G.E., Bateman A.;
 RT "Isolation and sequence of the granulin precursor cDNA from human
 RT bone marrow reveals tandem cysteine-rich granulin domains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:1715-1719(1992).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.

RX MEDLINE=93038704; PubMed=1417868;
 RA Bhandari V., Bateman A.;
 RT "Structure and chromosomal location of the human granulin gene.";
 RL Biochem. Biophys. Res. Commun. 188:57-63(1992).
 [3]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RC MEDLINE=92317004; PubMed=1618805;
 RX Plowman G.D., Green I.M., Neubauer M.G., Buckley S.D., McDonald V.L.,
 RA Todaro G.J., Shoyab M.;
 RT "The epithelin precursor encodes two proteins with opposing activities
 on epithelial cell growth.";
 RL J. Biol. Chem. 267:13073-13078(1992).
 [4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RC Yu W., Gibbs R.A.;
 RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 [5]
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP TISSUE=Cervix, and Lung;
 RC Strausberg R.;
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 RN SEQUENCE OF 206-233; 281-336; 364-396 AND 442-447.
 RP TISSUE=Leukocyte;
 RC MEDLINE=91097544; PubMed=2268320;
 RX Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;
 RA "Granulins, a novel class of peptide from leukocytes";
 RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).
 [7]
 RN FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY
 PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.
 [8]
 RN FUNCTION: GRANULIN A PROMOTES PROLIFERATION OF THE EPITHELIAL
 CELL LINE A431 IN CULTURE WHILE GRANULIN B ACTS AS AN ANTAGONIST
 TO GRANULIN A, INHIBITING THE GROWTH.
 [9]
 RN ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
 produced by alternative splicing.
 [10]
 RN TISSUE SPECIFICITY: IN MYELOGENOUS LEUKEMIC CELL LINES OF
 PROMONOCYTIC, PROMYELOCYTIC, AND PROERYTHROID LINEAGE, IN
 FIBROBLASTS, AND VERY STRONGLY IN EPITHELIAL CELL LINES. PRESENT
 IN INFLAMMATORY CELLS AND BONE MARROW. HIGHEST LEVELS IN KIDNEY.
 [11]
 RN PTM: GRANULINS ARE DISULFIDE BRIDGED.
 [12]
 RN This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 [13]
 RN EMBL; M75161; AAA58617.1; ALT_SEQ.
 DR EMBL; X62320; CAA4196.1; -;
 DR EMBL; AF055008; AAC09359.1; -;
 DR EMBL; BC000324; AAH00324.1; -;
 DR EMBL; BC010577; AAH10577.1; -;
 DR PIR; A38118; GYHU.
 DR PIR; D36698; D36698.
 DR PIR; JCL1284; JCL1284.
 DR Genew; HGNC:4601; GRN.
 DR MIM; 138945; -;
 DR InterPro; IPR000118; Granulin.
 DR Pfam; PF00396; granulin; 7.
 DR SMART; SM00277; GRAN; 7.
 DR PROSITE; PS00799; GRANULINS; 7.
 KW Cytokine; Repeat; Glycoprotein; Signal; Alternative splicing;
 RN Polymorphism.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 593 ACROGRANIN.
 FT PEPTIDE 18 747 PARAGRANULIN.
 FT PEPTIDE 758 ?113 GRANULIN 1.
 FT PEPTIDE ?123 ?179 GRANULIN 2.
 FT PEPTIDE 206 261 GRANULIN 3.

FT PEPTIDE 281 336 GRANULIN 4.
 FT PEPTIDE 364 7417 GRANULIN 5.
 FT PEPTIDE 442 7496 GRANULIN 6.
 FT PEPTIDE 7518 7573 GRANULIN 7.
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 377 531 MISSING (IN ISOFORM 2).
 FT VARIANT 454 454 G -> Q.
 FT CONFLICT 219 219 /FTID=VAR_003445.
 FT CONFLICT 386 386 S -> H (IN REF. 6).
 SQ SEQUENCE 593 AA; 63473 MW; 4E402BDB16DE2819 CRC64;
 Query Match 94.0%; Score 331; DB 1; Length 593;
 Best Local Similarity 93.0%; Pred. No. 1.6e-26;
 Matches 53; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 IGCDQHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
 Db 442 IGCDQHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498
 RESULT 4
 GRN_CAVPO STANDARD; PRT; 591 AA.
 ID GRN_CAVPO STANDARD; PRT; 591 AA.
 AC P28797;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2;
 DE Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7]
 DE (Fragment).
 GN GRN.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-21.
 RC TISSUE=Testis;
 RX MEDLINE=93228994; PubMed=8471244;
 RA Baba T., Hoff H.B. III, Nemoto H., Lee H., Orth J., Arai Y.,
 RA Gerton G.L.;
 RT "Acrogranin, an acrosomal cysteine-rich glycoprotein, is the
 precursor of the growth-modulating peptides, granulins, and
 epithelins, and is expressed in somatic as well as male germ cells.";
 RL Mol. Reprod. Dev. 34:233-243(1993).
 CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY
 PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC EMBL; M86735; AAA37030.1; -;
 DR InterPro; IPR000118; Granulin.
 DR Pfam; PF00396; granulin; 7.
 DR SMART; SM00277; GRAN; 6.
 DR PROSITE; PS00799; GRANULINS; 6.
 KW Cytokine; Repeat; Signal; Glycoprotein.
 RN Non-TER 1 1
 FT SIGNAL <1 3
 FT CHAIN 4 591 ACROGRANIN.
 FT PEPTIDE ?53 798 GRANULIN 1.

```

FT PEPTIDE 2111 2164 GRANULIN 2.
FT PEPTIDE 206 261 GRANULIN 3.
FT PEPTIDE 281 335 GRANULIN 4.
FT PEPTIDE 2376 2415 GRANULIN 5.
FT PEPTIDE 439 7494 GRANULIN 6.
FT PEPTIDE 2518 2569 GRANULIN 7.
FT CARBOHYD 183 183 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 591 AA; 62586 MW; EE7C9FC8F21CB8A1 CRC64;

Query Match
Best Local Similarity 87.2%; Score 307; DB 1; Length 591;
Matches 48; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPCPSLKGWACCOLPHAVCCEDRHCCPAGYTCNVKARTCEK 56
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 VGCDQHASCPRVQTCPCPLGGGWACCOLPHAVCCEDRGHCCPAGYTCNVKARSCEK 494

RESULT 5
CYSL_LYCES
ID CYSL_LYCES STANDARD; PRT; 346 AA.
AC P20721;
DT 01-FEB-1991 (Rel. 17, Last Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Low-temperature-induced cysteine proteinase precursor (EC 3.4.22.-)
DE (Fragment).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VENT Cherry;
RA Schaffer M.A.; Fischer R.L.;
RT Identifies a thiolprotease in tomato.;
RL Plant Physiol. 87:431-436(1988).
CC -!- INDUCTION: IN RESPONSE TO LOW TEMPERATURE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21444; AAA66308.1; -
CC PIR; JA0159; JA0159.
CC HSP; P00785; 2ACT.
CC MEROPS; C01.029; -
CC InterPro; IPR000118; Granulin.
CC InterPro; IPR000668; Peptidase_C1.
CC InterPro; IPR000169; SHprot_acsite.
CC Pfam; PF00112; Peptidase_C1; 1.
CC Pfam; PF00396; granulin; 1.
CC ProDom; PD000158; Peptidase_C1; 1.
CC SMART; SM00277; GRAN; 1.
CC PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
CC PROSITE; PS00639; THIOI_PROTEASE_HIS; 1.
CC PROSITE; PS00640; THIOI_PROTEASE_ASN; FALSE_NEG.
KW Hydrolase; Thiol protease; Zymogen; Glycoprotein.
FT NON_TER 1 1
FT PROPEP <1 17 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 18 346 LOW-TEMPERATURE-INDUCED CYSTEINE
FT PROTEINASE.
FT ACT_SITE 42 42 BY SIMILARITY.
FT ACT_SITE 178 178 BY SIMILARITY.
FT ACT_SITE 198 198 BY SIMILARITY.
FT DISULFID 39 81 BY SIMILARITY.

```

```

FT DISULFID 73 114 BY SIMILARITY.
FT DISULFID 172 223 BY SIMILARITY.
FT CARBOHYD 215 215 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 346 AA; 37429 MW; D42AC73944010928 CRC64;

Query Match
Best Local Similarity 44.9%; Score 158; DB 1; Length 346;
Matches 28; Conservative 5; Mismatches 19; Indels 6; Gaps 2;

Qy 3 CDOHTSCPVGQTCPCPSLKGWACCOLPHAVCCEDRHCCPAGYTCNVKARTC 54
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CDEYSQCAVGTCCTCCILQFRKSCFSWGCCPLEGATCCEDHYSCPHDPICNVROGTC 313

RESULT 6
ORYB_ORYSA
ID ORYB_ORYSA STANDARD; PRT; 471 AA.
AC P25777;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oryza beta chain precursor (EC 3.4.22.-)
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare; TISSUE=Seed;
RA MEDLINE=91358494; PubMed=1885617;
RA Watanabe H., Abe K., Emori Y., Hosoyama H., Arai S.;
RT "Molecular cloning and gibberellin-induced expression of multiple
RT cysteine proteinases of rice seeds (oryzains).";
RL J. Biol. Chem. 266:16897-16902(1991).
CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN SEEDS.
CC -!- INDUCTION: BY GIBBERELIC ACID (GA).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90407; BAA14403.1; -
CC PIR; JU0389; KHRZOB.
CC HSP; P00785; 2ACT.
CC MEROPS; C01.029; -
CC InterPro; IPR000118; Granulin.
CC InterPro; IPR000668; Peptidase_C1.
CC InterPro; IPR000169; SHprot_acsite.
CC Pfam; PF00112; Peptidase_C1; 1.
CC Pfam; PF00396; granulin; 1.
CC PRINTS; PR00705; PAPAIN.
CC ProDom; PD000158; Peptidase_C1; 1.
CC SMART; SM00277; GRAN; 1.
CC PROSITE; PS00640; THIOI_PROTEASE_ASN; 1.
CC PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
CC PROSITE; PS00639; THIOI_PROTEASE_HIS; 1.
KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 139 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 140 471 ORYZAIN BETA CHAIN.
FT ACT_SITE 164 164 BY SIMILARITY.
FT ACT_SITE 301 301 BY SIMILARITY.
FT ACT_SITE 321 321 BY SIMILARITY.
FT DISULFID 161 204 BY SIMILARITY.
FT DISULFID 195 237 BY SIMILARITY.
FT DISULFID 295 346 BY SIMILARITY.
FT CARBOHYD 340 340 N-LINKED (GLCNAC...) (POTENTIAL).

```

```

FT CARBOHYD 388 388 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 471 AA; 5505 MW; 0BF39D33995CEBID CRC64;

Query Match 42.6%; Score 150; DB 1; Length 471;
Best Local Similarity 44.8%; Pred. No. 2e-08;
Matches 26; Conservative 5; Mismatches 21; Indels 6; Gaps 2;

Qy 3 CDQHTSPVGQTCC-----PSLKGSWACCQLPHAVCCEDROHCCPGAGYTCNVKARTC 54
Db 385 CDDNFSPAGSTGCCAFGRNLCLVWGCPVEGATCKDKHASCCPPDPVCNTRAGTC 442
||| ||| ||| : | ||| : ||| : ||| ||| ||| : |||
||| ||| ||| : | ||| : ||| : ||| ||| ||| : |||

RESULT 7
GRN1_CYPCA STANDARD; PRT; 57 AA.
ID GRN1_CYPCA AC P81013;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Granulin 1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxId=7962;
RN [1]
SEQUENCE.
RC TISSUP=Spleen, and Kidney;
RX MEDLINE=93252781; PubMed=8486624;
RA Belcourt D.R., Lazure C., Bennett H.P.;
RT "Isolation and primary structure of the three major forms of
granulin-like peptides from hematopoietic tissues of a teleost fish
(Rt (Cyprinus carpio).";
RL J. Biol. Chem. 268:9230-9237(1993).
CC CC -I- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY.. THEY MAY
PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.
CC -I- TISSUE SPECIFICITY: UBIQUITOUS.
CC -I- PTM: GRANULINS ARE DISULFIDE BRIDGED.
DR InterPro; IPR000118; Granulin.
DR Pfam; PF00396; granulin; 1.
DR SMART; SM00277; GRAN; 1.
KW Cytokine.
SQ SEQUENCE 57 AA; 6289 MW; E4A131B1288FE55A CRC64;

Query Match 41.2%; Score 145; DB 1; Length 57;
Best Local Similarity 46.3%; Pred. No. 1.3e-08;
Matches 25; Conservative 2; Mismatches 27; Indels 0; Gaps 0;

Qy 1 IGCDHQHTSCPVGQTCCPSLKGSWACCQLPHAVCCEDROHCCPGAGYTCNVKARTC 54
Db 2 IHCDAAITICPDGTTCCLSPGYWVYCCPFSGQCQRDGIHCRRGHYCDSTSHC 55
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
GRN3_CYPCA STANDARD; PRT; 57 AA.
ID GRN3_CYPCA AC P81015;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Granulin 3.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxId=7962;
RN [1]
SEQUENCE.
RC TISSUP=Kidney;
RX MEDLINE=93252781; PubMed=8486624;
RA Belcourt D.R., Lazure C., Bennett H.P.;
RT "Isolation and primary structure of the three major forms of
granulin-like peptides from hematopoietic tissues of a teleost fish

```

OC Ehrhartoideae; Oryzeae; Oryza.
 RN NCBI_TaxID=4530;
 RC SEQUENCE FROM N.A., AND SEQUENCE OF 129-135.
 RX STRAIN=cv. Nipponbare; TISSUE=Seed;
 RM MEDLINE=91358494; PubMed=1885617;
 RA Watanabe H., Abe K., Emori Y., Hosoyama H., Arai S.;
 RT "Molecular cloning and gibberellin-induced expression of multiple
 cysteine proteinases of rice seeds (oryzains).";
 RL J. Biol. Chem. 266:16897-16902(1991).
 CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN SEEDS.
 CC -1- INDUCTION: BY GIBBERELIC ACID (GA).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90406; BAA14402.1;
 DR PIR; J03388; KHRZOA.
 DR HSP; P00785; 2ACT.
 DR MEMOPS; C01.029; -.
 DR InterPro; IPR000118; Granulin.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprot_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR Pfam; PF00396; granulin; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR ProDom; PD000158; Peptidase_C1; 1.
 DR SMART; SM00277; GRAN; 1.
 DR PROSITE; PS00139; THIOLE PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOLE PROTEASE_HIS; 1.
 DR PROSITE; PS00640; THIOLE PROTEASE_ASN; 1.
 DR Hydrolase; thiol protease; Zymogen; Glycoprotein; Signal.
 KW SIGNAL
 FT CHAIN 1 21 POTENTIAL.
 FT PROPEP 22 128 ACTIVATION PEPTIDE.
 FT CHAIN 129 458 ORYZAIN ALPHA CHAIN.
 FT ACT_SITE 153 153 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.
 FT ACT_SITE 309 309 BY SIMILARITY.
 FT DISULFID 150 192 BY SIMILARITY.
 FT DISULFID 184 225 BY SIMILARITY.
 FT DISULFID 283 334 BY SIMILARITY.
 FT CARBOHYD 445 445 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 458 AA; 50274 MW; BIA0907750735CC7 CRC64;
 Query Match 38.6%; Score 136; DB 1; Length 458;
 Best Local Similarity 41.4%; Pred. No. 4.9e-07;
 Matches 24; Conservative 6; Mismatches 22; Indels 6; Gaps 2;
 QY 3 CDQHTSPVGTCCPSLK-----GSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTC 54
 DB 367 CDNYITCPDSTTCCCIYEYKCYVANGCCPLEGATCCDDHSCCPHEYPICNVQGGTC 424
 RESULT 11
 PMDL1_L0CMI STANDARD; PRT; 54 AA.
 ID PMDL1_L0CMI
 AC P80059;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE Pars intercerebralis major peptide D1 (PMP-D1).
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
 OC Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
 OX NCBI_TaxID=7004;
 RN [1]

RP SEQUENCE.
 RC TISSUE=Pars intercerebralis;
 RX MEDLINE=92155197; PubMed=1740125;
 RA Nakakura N., Hietter H., van Dorselaer A., Luu B.;
 RT "Isolation and structural determination of three peptides from the
 insect Locusta migratoria. Identification of a deoxyhexose-linked
 peptide.";
 RL Eur. J. Biochem. 204:147-153(1992).
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- PTM: SIX DISULFIDE BONDS ARE PRESENT.
 CC -1- SIMILARITY: STRONG, TO GRANULINS.
 DR PIR; S23075; S23075.
 DR InterPro; IPR000118; Granulin.
 DR Pfam; PF00396; granulin; 1.
 DR SMART; SM00277; GRAN; 1.
 DR PROSITE; PS00799; GRANULINS; 1.
 KW Neuropeptide.
 SQ SEQUENCE 54 AA; 5788 MW; 3011D739060D72F5 CRC64;
 Query Match 37.4%; Score 131.5; DB 1; Length 54;
 Best Local Similarity 42.3%; Pred. No. 2.8e-07;
 Matches 22; Conservative 6; Mismatches 23; Indels 1; Gaps 1;
 QY 3 CDQHTSPVGTCCPSLKSGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTC 54
 DB 2 CTEXT-CPGTETCTTPQEGEGCCPYKEGVCCLDGIHCCSPGTVCDEDHRC 52
 RESULT 12
 RD21_ARATH STANDARD; PRT; 462 AA.
 ID RD21_ARATH
 AC P43297;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteine proteinase RD21A precursor (EC 3.4.22.-).
 GN RD21A OR ATIG47128 OR F2G19.31.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RM MEDLINE=93314960; PubMed=8325504;
 RX Koizumi M., Yamaguchi-Shinozaki K., Tsuji H., Shinozaki K.;
 RA "Structure and expression of two genes that encode distinct drought-
 inducible cysteine proteinases in Arabidopsis thaliana.";
 RT Gene 129:175-182(1993).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RM MEDLINE=21016719; PubMed=11130712;
 RX Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 Buehler E.K., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzlahti A.,
 Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 Pai G., Peterson J., Pham P.K., Rizso M., Rooney T., Rowley D.,
 Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 Uteckback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).


```
CC -!- INDUCTION: BY HIGH SALT CONDITIONS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13043; BAA02374.1; -.
DR EMBL; AC083835; AAG50628.1; -.
DR HSSP; P00785; 2ACT.
DR MEROPS; C01.029; -.
DR InterPro; IPR000118; Granulin.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR Pfam; PF00396; granulin; 1.
DR PRINTS; PR00705; PAPAIN.
DR ProDom; PD00158; Peptidase_C1; 1.
DR SMART; SM00277; GRAN; 1.
DR SMART; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
DR Hydrolase; Thiol protease; zymogen; Glycoprotein; Signal.
KW SIGNAL 1 21
FT PROPEP 22 136 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 137 462 CYSTEINE PROTEINASE RD21A.
FT ACT_SITE 161 161 BY SIMILARITY.
FT ACT_SITE 297 297 BY SIMILARITY.
FT ACT_SITE 317 317 BY SIMILARITY.
FT DISULFID 158 200 BY SIMILARITY.
FT DISULFID 192 233 BY SIMILARITY.
FT DISULFID 291 342 BY SIMILARITY.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 462 AA; 50966 MW; 4C59955CCB95AA58 CRC64;

Query Match 36.9%; Score 130; DB 1; Length 462;
Best Local Similarity 37.9%; Pred. No. 2e-06; 23; Indels 6; Gaps 2;
Matches 22; Conservative 7; Mismatches 23;

QY 3 CDQHTSCPVGQTCPSLKGWACCOLPHAVCCEDRQHCPCAGY-TCNVKARTC 54
DB 375 CDSYVTCPSNTCCCLFEYKGYCFWAGCCPLEATCCDDNYSCCPHEYPVCDLDQGTG 432

RESULT 13
ENAL_HORSE
ID ENAL_HORSE STANDARD; PRT; 46 AA.
AC P80930;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Antimicrobial peptide enAP-1 (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE.
RX MEDLINE=92347572; PubMed=1639474;
RA Couto A.M., Harwig S.S.L., Cullor J.S., Hughes J.P., Lehrer R.I.;
RT "Identification of enAP-1, an antimicrobial peptide from equine
neutrophils."
RL Infect. Immun. 60:3065-3071(1992).
CC -!- FUNCTION: HAS ANTIMICROBIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC GRAM-POSITIVE BACTERIA.
CC -!- SIMILARITY: SEEMS TO CORRESPONDS TO GRANULIN 7.
DR InterPro; IPR000118; Granulin.
```

```
DR Pfam; PF00396; granulin; 1.
DR SMART; SM00277; GRAN; 1.
DR PROSITE; PS00799; GRANULINS; PARTIAL.
KW Antibiotic. 46
FT NON_TER 46
SQ SEQUENCE 46 AA; 4888 MW; 2171934C15265862 CRC64;

Query Match 36.4%; Score 128; DB 1; Length 46;
Best Local Similarity 42.2%; Pred. No. 5.5e-07;
Matches 19; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 TCGDQHTSCPVGQTCPSLKGWACCOLPHAVCCEDRQHCPCAGY 45
DB 2 VOCGEGHFCHDXQTCCRASQGGXACCPYSQGVCCADQRHCCPVGF 46

RESULT 14
MCS_HUMAN
ID MCS_HUMAN STANDARD; PRT; 116 AA.
AC P49901; Q96A42;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm mitochondrial capsule selenoprotein (MCS).
GN MCSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96429994; PubMed=8833144;
RA Aho H., Schwemmer M., Tessmann D., Murphy D., Mattei M.-G., Engel W.,
RA Adham I.M.;
RT "Isolation, expression, and chromosomal localization of the human
mitochondrial capsule selenoprotein gene (MCSP).";
RL Genomics 32:184-190(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STRUCTURAL PROTEIN OF THE SPERM MITOCHONDRIAL CAPSULE.
IMPORTANCE OF THE MAINTENANCE AND STABILIZATION OF THE CRESCENT
STRUCTURE OF THE SPERM MITOCHONDRIA.
CC -!- TISSUE SPECIFICITY: TESTIS.
CC -!- DEVELOPMENTAL STAGE: POSTMEIOTIC STAGES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X89960; CAA62000.1; -.
DR EMBL; BC014593; AAH14593.1; -.
DR EMBL; BC016744; AAH16744.1; -.
DR Genew; HGNC:6962; MCSP.
DR MIM; 601148; -.
RW Mitochondrion; Selenium; Sperm; Testis; Spermatogenesis; Repeat.
FT DOMAIN 6 68 7 x 7 (OR 8) AA APPROXIMATE REPEATS.
FT REPEAT 6 13 1.
FT REPEAT 14 21 2.
FT REPEAT 30 37 3.
FT REPEAT 38 45 4.
FT REPEAT 46 53 5.
FT REPEAT 54 61 6.
FT REPEAT 62 68 7.
FT REPEAT 63 63 P -> A (IN REF. 1).
SQ SEQUENCE 116 AA; 12767 MW; 242191C60AEE03B2 CRC64;
```


DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00054; laminin_G; 1.
DR Pfam; PF00594; gla; 1.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00025; LAM_G_DOMAIN; 2.
KW Plasma; Gamma-carboxyglutamic acid; Calcium; Vitamin K; Zymogen;
KW Blood coagulation; Hydroxylation; Glycoprotein; Signal; Repeat;
KW EGF-like domain; Polymorphism; Disease mutation; Thrombophilia.
FT SIGNAL 1 24
FT PROPEP 25 41
FT CHAIN 42 676
FT DOMAIN 42 86
FT DOMAIN 87 116
FT DOMAIN 117 155
FT DOMAIN 157 200
FT DOMAIN 201 242
FT DOMAIN 243 283
FT DOMAIN 299 475
FT DOMAIN 484 666
FT MOD_RES 47 47
FT MOD_RES 48 48
FT MOD_RES 55 55
FT MOD_RES 57 57
FT MOD_RES 60 60
FT MOD_RES 61 61
FT MOD_RES 66 66
FT MOD_RES 67 67
FT MOD_RES 70 70
FT MOD_RES 73 73
FT MOD_RES 77 77
FT MOD_RES 136 136
FT DISULFID 121 134
FT DISULFID 126 143
FT DISULFID 145 154
FT DISULFID 161 175
FT DISULFID 171 184
FT DISULFID 186 199
FT DISULFID 205 217
FT DISULFID 212 226
FT DISULFID 228 241
FT DISULFID 247 256
FT DISULFID 252 265
FT DISULFID 267 282
FT CARBOHYD 499 499
FT CARBOHYD 509 509
FT CARBOHYD 530 530
FT VARIANT 196 196
FT VARIANT 258 258
FT VARIANT 501 501
FT CONFLICT 11 11
FT CONFLICT 26 26
SQ SEQUENCE 676 AA; 75122 MW; 2B88A04F85403F25 CRC64;

Query Match 24.0%; Score 84.5; DB 1; Length 676;
Best Local Similarity 30.8%; Pred. No. 0.095;

Matches 24; Conservative 8; Mismatches 17; Indels 29; Gaps 5;
Qy 2 GCDQ-----HTSCPVG-----OTC-----C---PSLKGSWACCOLPHAVCCEDR 37
Db 170 GCSQICDNTPGSYHCSCKNGFVNLNKKDKDVEDCSLKPSICGTAVCKNIPGDFECE-- 227
Qy 38 QHCCPAGYTCNVKARTCE 55
Db 228 ---CPEGRYNLKSKSCE 242

Search completed: April 21, 2003, 14:21:13
Job time : 28 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:17:50 ; Search time 84 Seconds
(without alignments)
139.818 Million cell updates/sec

Title: US-09-874-056-4
Perfect score: 352
Sequence: 1 IGCQHTSCPVGQTCPSLK.....QHCCPAGYTCNVKARTCEKQ 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	347	98.6	589	11 Q9D2V3	Q9d2v3 mus musculus
2	338	96.0	413	4 Q9H8S1	Q9h8s1 homo sapien
3	198	53.4	621	13 Q8CGN9	Q8cgn9 brachydanio
4	176	50.0	358	5 Q9U362	Q9u362 caenorhabdi
5	168	47.7	466	10 Q9ST61	Q9st61 solanum tub
6	158	44.9	466	10 Q49877	Q49877 lycopersico
7	157	44.6	147	13 Q90ZC9	Q90zc9 brachydanio
8	157	44.6	147	13 Q90ZC8	Q90zc8 brachydanio
9	152	43.2	147	13 Q90ZD0	Q90zd0 brachydanio
10	152	43.2	445	10 Q8W182	Q8w182 brassica ol
11	150	42.6	122	10 Q854X9	Q854x9 oryza sativ
12	147	41.8	57	13 Q9PRN7	Q9prn7 carassius a
13	143	40.6	452	10 Q9LT78	Q9lt78 arabidopsis
14	142	40.3	427	10 Q43423	Q43423 dianthus ca
15	141	40.1	437	10 Q93XC2	Q93xc2 arabidopsis
16	141	40.1	463	10 Q9FMH8	Q9fmh8 arabidopsis

17	141	40.1	465	10 Q9SLY9	Q9sly9 zea mays (m
18	140	39.8	416	10 Q04507	Q04507 arabidopsis
19	140	39.8	454	10 Q04922	Q04922 pseudotsuga
20	139	38.5	493	10 Q22499	Q22499 zea mays (m
21	136	38.6	370	10 Q9XF79	Q9xf79 sandersonia
22	136	38.6	461	10 Q9FSS0	Q9fss0 oryza sativ
23	136	38.6	461	10 Q8W181	Q8w181 brassica ol
24	132	37.5	462	10 Q93XQ9	Q93xq9 ipomoea bat
25	132	37.5	480	10 Q22500	Q22500 zea mays (m
26	130	36.9	129	10 Q9SAM9	Q9sam9 arabidopsis
27	130	36.9	455	10 Q24323	Q24323 phaseolus v
28	130	36.9	462	10 Q94BX1	Q94bx1 arabidopsis
29	128	36.4	485	10 Q8W180	Q8w180 brassica ol
30	125.5	35.7	464	10 Q41064	Q41064 pisum sativ
31	117.5	33.4	210	10 Q9SP97	Q9sp97 solanum cha
32	103	29.3	501	10 Q9SP93	Q9sp93 matricaria
33	88.5	25.1	966	5 Q22378	Q22378 caenorhabdi
34	88	25.0	1280	11 Q9EPX2	Q9epx2 mus musculu
35	83	23.6	162	5 Q8WSW3	Q8wsW3 tetrahymena
36	81.5	23.2	615	13 Q57409	Q57409 brachydanio
37	81.5	23.2	642	13 P79941	P79941 xenopus lae
38	81	23.0	422	10 Q9FVR4	Q9fvr4 arabidopsis
39	81	23.0	730	5 Q45021	Q45021 caenorhabdi
40	80.5	22.9	389	5 Q21081	Q21081 caenorhabdi
41	80.5	22.9	686	11 Q9DBU9	Q9dbu9 mus musculu
42	80.5	22.9	721	13 Q91902	Q91902 xenopus lae
43	80.5	22.9	723	4 Q9UJV2	Q9ujv2 homo sapien
44	80.5	22.9	723	4 Q9NU41	Q9nu41 homo sapien
45	80.5	22.9	728	13 Q90656	Q90656 gallus gall

ALIGNMENTS

RESULT 1

Q9D2V3	PRELIMINARY;	PRT;	589 AA.
ID	Q9D2V3		
AC	Q9D2V3		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	Adult male kidney cDNA, RIKEN full-length enriched library,		
DE	clone:0610012H06, full insert sequence.		
GN	GRN.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,		
RA	Hayashizaki Y.		
RT	Functional annotation of a full-length mouse cDNA collection."		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK018744; BAB31384.1;		


```

RN SEQUENCE FROM N.A.
RC STRAIN=CV. STIRLING; TISSUE=LEAF;
RX MEDLINE=20089145; PubMed=10624019;
RA AYOVA A.O., Stewart H.E., de Jong W., Heilbronn J., Lyon G.D.,
RA Birch P.R.J.;
RT "A cysteine protease gene is expressed early in resistant potato
RT interactions with Phytophthora infestans.";
RL Mol. Plant Microbe Interact. 12:1114-1119(1999).
DR EMBL: AJ245924; CAB53515.1; -.
DR HSP: P00785; 2ACT.
DR MEROPS: C01.029; -.
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR001211; PhospholipaseA2.
DR InterPro: IPR000169; SHprot_acsite.
DR Pfam: PF00396; granulin; 1.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.
DR ProDom: PD000158; Peptidase_C1; 1.
DR SMART: SM00277; GRAN; 1.
DR PROSITE: PS00118; PA2_HIS; UNKNOWN_1.
DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1.
DR PROSITE: PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Hydrolyase; Protease; Thiol protease.
SQ SEQUENCE 466 AA; 50919 MW; A6BB2A406F9A2DA1 CRC64;

Query Match 47.7%; Score 168; DB 10; Length 466;
Best Local Similarity 50.0%; Pred. No. 1.le-14;
Matches 29; Conservative 5; Mismatches 18; Indels 6; Gaps 2;

QY 3 CDQHTSCPVGQTCPCPSLK-----GSWACQQLPHAVCCEDRQHCCPAGY-TCNVKARTC 54
   ||:: |||| ||| |:: || || | |||| ||| ||| ||| |
Db 376 CDEYSQCPVGTTCVCCVLEFRRCFSFGCCPLEGATCCEDHSCCPHDYPCVNCVRQGT 433

RESULT 6
O49877
ID O49877 PRELIMINARY; PRT; 466 AA.
AC O49877;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CYPL.
DE C14 OR TDI-65.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CHERRY;
RA Lers A., Burd S., Sonogo L., Khalchitski A., Lomanec E.;
RL Plant Physiol. 116:1193-1193(1998).
RN [2]
RA Harrak H., Tabaeizadeh Z.;
RT "Isolation and characterization of a drought-induced gene (tdi-65)
RT encoding a cysteine protease in Lycopersicon esculentum.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ003137; CAA05894.1; -.
DR EMBL: AF172856; AAD48496.1; -.
DR HSP: P00785; 2ACT.
DR MEROPS: C01.029; -.
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR001211; PhospholipaseA2.
DR InterPro: IPR000169; SHprot_acsite.
DR Pfam: PF00396; granulin; 1.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.

```

```

DR ProDom: PD000158; Peptidase_C1; 1.
DR SMART: SM00277; GRAN; 1.
DR PROSITE: PS00118; PA2_HIS; UNKNOWN_1.
DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1.
DR PROSITE: PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Hydrolyase; Protease; Thiol protease.
SQ SEQUENCE 466 AA; 51141 MW; FELE468967508BC7 CRC64;

Query Match 44.9%; Score 158; DB 10; Length 466;
Best Local Similarity 48.3%; Pred. No. 2.4e-13;
Matches 28; Conservative 5; Mismatches 19; Indels 6; Gaps 2;

QY 3 CDQHTSCPVGQTCPCPSLK-----GSWACQQLPHAVCCEDRQHCCPAGY-TCNVKARTC 54
   ||:: |||| ||| |:: || || | |||| ||| ||| ||| |
Db 376 CDEYSQCAVGTTCVCCILQFRRCFSFGCCPLEGATCCEDHSCCPHDYPCVNCVRQGT 433

RESULT 7
Q90ZC9
ID Q90ZC9 PRELIMINARY; PRT; 147 AA.
AC Q90ZC9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Granulin 2.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Cadieux B., Bennett H.P.J.;
RT "Zebrafish granulin 1, 2, and hybrid cDNA sequences.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF273480; AAK58709.1; -.
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR002221; WAP.
DR Pfam: PF00396; granulin; 1.
DR PROSITE: PS00317; 4_DISULFIDE_CORE; UNKNOWN_1.
DR PROSITE: PS00317; 4_DISULFIDE_CORE; UNKNOWN_1.
SQ SEQUENCE 147 AA; 16434 MW; 2A4B4A67D0D1DEB6 CRC64;

Query Match 44.6%; Score 157; DB 13; Length 147;
Best Local Similarity 46.3%; Pred. No. 1.2e-13;
Matches 25; Conservative 2; Mismatches 27; Indels 0; Gaps 0;

QY 1 ICDDQHTSCPVGQTCPCPSLKGSWACQQLPHAVCCEDRQHCCPAGYTCNVKARTC 54
   | | | | | | | | | | | | | | | | | | | | | | |
Db 36 IHCDAVTGCPDRTTCRTPYGKWTCCPPMGQCCRDGIHCCRHGRYRCNFASTRC 89

RESULT 8
Q90ZC8
ID Q90ZC8 PRELIMINARY; PRT; 147 AA.
AC Q90ZC8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hybrid granulin.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Cadieux B., Bennett H.P.J.;
RT "Zebrafish granulin 1, 2, and hybrid cDNA sequences.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF273481; AAK58710.1; -.
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR002221; WAP.

```


DR Pfam; PF00396; granulin; 1.
DR SMART; SM00277; GRAN; 1.
SQ SEQUENCE 57 AA; 6321 MW; EAAL2E4EC8901ABA CRC64;

Query Match 41.8%; Score 147; DB 13; Length 57;
Best Local Similarity 46.3%; Pred. No. 1.le-12;
Matches 25; Conservative 2; Mismatches 27; Indels 0; Gaps 0;

QY 1 IGCQHHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCCPAGYTCTNVKARTC 54
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 2 IHCDSSITICPDGTTCCLSPGVVWVCPPFSMGQCRCRDGIHCCRHRGYHCDSTSTHC 55

RESULT 13
Q9LT78 ID Q9LT78 PRELIMINARY; PRT; 452 AA.

AC Q9LT78;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cysteine proteinase.
GN AT3G19400, MLD14.12 OR AT3G19400.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3703;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=COLUMBIA;
RC MEDLINE=20277480; PubMed=10819329;
RX Nakamura Y.;
RA "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty pI and TAC clones.";
RT DNA Res. 7:131-135(2000).
[3]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
[4]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[5]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[6]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[7]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[8]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[9]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[10]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[11]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[12]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[13]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[14]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[15]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[16]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[17]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[18]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[19]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[20]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[21]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawal J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
[22]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Che

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=XYLEM, AND BARK;
 RA Funk V., Kositsup B., Zhao C., Beers E.P.;
 RT "The Arabidopsis xylem peptidase XCP1 is a tracheary element vacuolar
 RT protein that may be a papain ortholog.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF388175; AAK71314.1; -
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR000118; Granulin.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprot_acsite.
 DR InterPro; IPR000834; Zn_carbopept.
 DR Pfam; PF00396; granulin; 1.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR ProDom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CVS; UNKNOWN_1.
 DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; UNKNOWN_1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; UNKNOWN_1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 SQ SEQUENCE 437 AA; 48142 MW; 2CEB94E6E340DA5E CRC64;

Query Match 40.1%; Score 141; DB 10; Length 437;
 Best Local Similarity 41.0%; Pred. No. 4.5e-11;
 Matches 25; Conservative 9; Mismatches 21; Indels 6; Gaps 2;
 QY 3 CDQHTCPVGTCCPS-----LKGSWACCOLPHAVCCEDRQHCPCAGY-TCNVKARTCEK 56
 Db 350 CNLFYCSSGETCCARELFGCLCFSWKCCCEIESAVCCCKGRHCCPHDYPVCDTTRSLCLK 409
 QY 57 Q 57
 Db 410 K 410

Search completed: April 21, 2003, 14:22:45
 Job time : 85 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:19:15 ; Search time 44 Seconds
(without alignments)
124.538 Million cell updates/sec

Title: US-09-874-056-4

Perfect score: 352

Sequence: 1 IGCQHTSCPVGQTCPCSLK.....QHCCPAGYTCNVKARTCEKQ 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	347	98.6	589	2 C38128	epithelin/granulin
2	347	98.6	589	2 B38128	epithelin/granulin
3	338	96.0	593	1 GHU	granulin precursor
4	307	87.2	591	2 I48141	acroganin - guine
5	176	50.0	318	2 E87929	protein T22H2.6 [i
6	176	50.0	345	2 T25138	hypothetical prote
7	176	50.0	358	2 T25137	hypothetical prote
8	158	44.9	346	2 JA0159	cysteine proteins
9	158	44.9	466	2 T08416	cysteine proteins
10	150	42.6	471	1 KHR20B	oryzain (EC 3.4.22
11	145	41.2	57	2 A46654	growth modulatory
12	144	40.9	57	2 C46654	growth modulatory
13	142	40.3	427	2 S57776	cysteine proteins
14	140	39.8	57	2 B46654	growth modulatory
15	140	39.8	416	2 G86232	cysteine proteins
16	140	39.8	454	2 JC4848	cysteine proteins
17	139	39.5	493	2 T01206	cysteine proteins
18	136	38.6	458	1 KHR20A	oryzain (EC 3.4.22
19	132	37.5	480	2 T01207	cysteine proteins
20	131.5	37.4	54	1 S23075	protein PMP-DI - m
21	130	36.9	455	2 T12041	cysteine proteins
22	130	36.9	462	2 JN0719	drought-inducible
23	128	36.4	46	2 A44794	antimicrobial pept
24	125.5	35.7	464	2 S24602	cysteine proteins
25	88.5	25.1	1101	2 T16840	hypothetical prote
26	84.5	24.0	676	1 KXHUS	plasma protein S p
27	82.5	23.4	642	2 S53434	plasma protein S p
28	82.5	23.4	685	2 JC7570	Delta-4 protein -
29	82.5	23.4	722	2 I48324	DELTA-like 1 - mou

30	81	23.0	61	2 JC1449	metallothionein A
31	81	23.0	422	2 D86446	hypothetical prote
32	80.5	22.9	389	2 T23167	hypothetical prote
33	80.5	22.9	686	2 JC7569	Delta-4 protein -
34	80.5	22.9	728	2 T50719	C-Delta-1 - chicke
35	79	22.4	4660	2 T42737	gp330 protein prec
36	78	22.2	61	1 SMOB02	metallothionein II
37	78	22.2	1820	2 A55494	metallothionein Ia
38	77.5	22.0	435	2 S40993	latent transformin
39	77.5	22.0	1391	2 T20406	hypothetical prote
40	77.5	22.0	61	2 I46602	metallothionein -
41	77	21.9	63	2 C34620	metallothionein -
42	77	21.9	63	2 S33381	metallothionein -
43	77	21.9	63	2 A34620	metallothionein -
44	77	21.9	63	2 A34620	metallothionein -
45	77	21.9	63	2 A34958	metallothionein -

ALIGNMENTS

RESULT 1

C38128

epithelin/granulin precursor - mouse

N:Alternate names: acroganin; PC-cell-derived growth factor

C:Species: Mus musculus (house mouse)

C>Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 20-Aug-1999

C:Accession: C38128; S32503; I49468; A46705

R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, J. Biol. Chem. 267, 13073-13078, 1992

A:Title: The epithelin precursor encodes two proteins with opposing activities on epi

A:Reference number: A38128; MUID:92317004; PMID:1618805

A:Accession: C38128

A:Molecule type: mRNA

A:Residues: 1-589 <P>

A:Cross-references: GB:62321; MID:950851; PIDN:CAA44197.1; PID:950852

R:Baba, T.; Nemoto, H.; Watanabe, K.; Arai, Y.; Gerton, G.L.

FEBS Lett. 322, 89-94, 1993

A:Title: Exon/Intron organization of the gene encoding the mouse epithelin/granulin p

A:Reference number: S32503; MUID:93245991; PMID:8482392

A:Accession: S32503

A:Molecule type: DNA

A:Residues: 18-349, 'L', 351-589 <BAB>

R:Baba, T.; Hoff, H.B.

Mol. Reprod. Dev. 34, 233-243, 1993

A:Title: Acroganin, an acrosomal cysteine-rich glycoprotein, is the precursor of the

A:Reference number: I48141; MUID:93228994; PMID:8471244

A:Accession: I49468

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-250, 'L', 252-253, 'V', 255-349, 'L', 351-401, 'SA', 404-589 <RES>

A:Cross-references: GB:M86736; MID:g191766; PIDN:AAA37191.1; PID:g191767

R:Zhou, J.; Gao, G.; Crabb, J.W.; Serrero, G.

J. Biol. Chem. 268, 10863-10869, 1993

A:Title: Purification of an autocrine growth factor homologous with mouse epithelin p

A:Reference number: A46705; MUID:93266526; PMID:8496151

A:Accession: A46705

A>Status: preliminary

A:Molecule type: protein

A:Residues: 18-19, 'X', 21-25, 'X', 27-29, 'XX', 32, 'XXX', 119-127, 152-154, 'DXK', 158-161, 'X'

C:Superfamily: granulin

Query Match 98.6%; Score 347; DB 2; Length 589;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGQTCPCSLKSGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56

|||||

Db 440 IGCQHTSCPVGQTCPCSLKSGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 2

B38128

epithelin/granulin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 20-Aug-1999
C:Accession: B38128; A36199; B36199; E36698; I53272
R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.
J. Biol. Chem. 267, 13073-13078, 1992
A:Title: The epithelin precursor encodes two proteins with opposing activities on epithelial cells
A:Reference number: A38128; MUID:92317004; PMID:1618805
A:Accession: B38128
A:Molecule type: mRNA
A:Residues: 1-589 <PLO>
A:Cross-references: GB:M62322; NID:g56108; PIDN:CAA4198.1; PID:g56109
R:Shoyab, M.; McDonald, V.L.; Bytes, C.; Todaro, G.J.; Plowman, G.D.
Proc. Natl. Acad. Sci. U.S.A. 87, 7912-7916, 1990
A:Title: Epithelins 1 and 2: Isolation and characterization of two cysteine-rich growth factors
A:Reference number: A36199; MUID:91045907; PMID:2236009
A:Accession: A36199
A:Molecule type: protein
A:Residues: 280-300 <SHO>
A:Accession: B36199
A:Molecule type: protein
A:Residues: 205-226 <SH2>
R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.
Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990
A:Title: Granulins, a novel class of peptide from leukocytes.
A:Reference number: A36698; MUID:91097544; PMID:2268320
A:Accession: E36698
A:Molecule type: protein
A:Residues: 279-307, 'SB', 310-324, 'T', 326, 'X', 328, 'Q' <BAT>
R:Bhandari, V.; Glaid, A.; Bateman, A.
Endocrinology 133, 2682-2689, 1993
A:Title: The complementary deoxyribonucleic acid sequence, tissue distribution, and cell expression of the human granulins
A:Reference number: I53272; MUID:94062640; PMID:8243292
A:Accession: I53272
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-200, 'S', 203-388, 'M', 390-589 <RES>
A:Cross-references: GB:M97750; NID:g204223; PIDN:AAA16903.1; PID:g204224
C:Superfamily: granulin

Query Match 98.6%; Score 347; DB 2; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495
|||||

RESULT 3
GIHU
N:Alternate names: [validated] - human
N:Contains: granulin A; granulin B; granulin C; granulin D; granulin E; granulin F; granulin G
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1992 #sequence_revision 03-May-1996 #text_change 08-Dec-2000
C:Accession: JCI284; A38128; A38118; A36698; B36698; D36698; A56873
R:Bhandari, V.; Bateman, A.
Biochem. Biophys. Res. Commun. 188, 57-63, 1992
A:Title: Structure and chromosomal location of the human granulin gene.
A:Reference number: JCI284; MUID:93038704; PMID:1417868
A:Accession: JCI284
A:Molecule type: DNA
A:Residues: 1-593 <BHA>
R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.
J. Biol. Chem. 267, 13073-13078, 1992
A:Title: The epithelin precursor encodes two proteins with opposing activities on epithelial cells
A:Reference number: A38128; MUID:92317004; PMID:1618805
A:Accession: A38128
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-593 <PLO>
A:Cross-references: GB:M62320; NID:g31192; PIDN:CAA44196.1; PID:g31193

R:Bhandari, V.; Palfree, R.G.E.; Bateman, A.
Proc. Natl. Acad. Sci. U.S.A. 89, 1715-1719, 1992
A:Title: Isolation and sequence of the granulin precursor cDNA from human bone marrow
A:Reference number: A38118; MUID:92179253; PMID:1542665
A:Accession: A38118
A:Molecule type: mRNA
A:Residues: 1-406, 'R', 408-433, 'G', 435-453, 'G', 455-459, 'Q', 461-546, 'A', 548-566, 'R', 568
A:Cross-references: GB:M75161; NID:g183612; PIDN:AAA58617.1; PID:g183613
A:Note: this sequence has been revised in reference JCI284
R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.
Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990
A:Title: Granulins, a novel class of peptide from leukocytes.
A:Reference number: A36698; MUID:91097544; PMID:2268320
A:Accession: A36698
A:Molecule type: protein
A:Residues: 281-336 <BAT>
A:Note: this protein was purified and characterized as granulin A
A:Accession: B36698
A:Molecule type: protein
A:Residues: 206-218, 'H', 220-233 <BA2>
A:Note: this protein was purified and characterized as granulin B
A:Accession: C36698
A:Molecule type: protein
A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>
A:Note: this protein was purified and characterized as granulin C
A:Accession: D36698
A:Molecule type: protein
A:Residues: 442-446, 'XDTSS', 456-458, 'DG', <BA4>
R:Kardana, A.; Bagshawe, K.D.; Coles, B.; Read, D.; Taylor, M.
Br. J. Cancer 67, 686-692, 1993
A:Title: Characterisation of DGP and its relationship with beta-core fragment.
A:Reference number: A56873; MUID:9329246; PMID:8471426
A:Accession: A56873
A:Molecule type: protein
A:Residues: 281-283, 'X', 285-289, 'S', 291-295 <KAR>
A:Experimental source: urine
A:Note: sequence extracted from NCBI backbone (NCBIP:129524)
C:Genetics:
A:Gene: GDB:GRN
A:Cross-references: GDB:136006; OMIM:138945
A:Map position: 17pter-17qter
A:Introns: 46/3; 88/3; 117/1; 154/3; 200/1; 236/3; 279/1; 311/3; 393/3; 471/3; 548/3
C:Superfamily: granulin
C:Keywords: glycoprotein; tandem repeat
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-593/Product: granulin #status predicted <MAT>
F:18-593/Product: progranulin #status predicted <PRO>
F:18-44/Product: paraganulin #status experimental <PGR>
F:58-113/Product: granulin G #status predicted <GRG>
F:123-179/Product: granulin F #status predicted <GRF>
F:206-261/Product: granulin B #status experimental <GRB>
F:281-336/Product: granulin A #status experimental <GRA>
F:364-417/Product: granulin C #status experimental <GRC>
F:442-496/Product: granulin D #status predicted <GRD>
F:518-573/Product: granulin E #status predicted <GRE>
F:368/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 96.0%; Score 338; DB 1; Length 593;
Best Local Similarity 94.7%; Pred. No. 1e-25;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 57
|||||
Db 442 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEK 498
|||||

RESULT 4
I48141
acroganin - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C:Accession: I48141
R:Baba, T.; Hoff, H.B.

RESULT 11
A46654
growth modulatory factor granulin-1 - common carp
C:Species: Cyprinus carpio (common carp)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 10-Nov-1995
C:Accession: A46654; A40180
R:Belcourt, D.R.; Lazure, C.; Bennett, H. P.
J. Biol. Chem. 268, 9230-9237, 1993
A:Title: Isolation and primary structure of the three major forms of granulin-like peptide
A:Reference number: A46654; MUID:93252781; PMID:8486624
A:Accession: A46654

RESULT 14
B46654

growth modulatory factor granulin-2 - common carp

Search completed: April 21, 2003, 14:23:35
Job time : 45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:20:45 ; Search time 16 Seconds
(without alignments)
269.374 Million cell updates/sec

Title: US-09-874-056-4
Perfect score: 352
Sequence: 1 IGCQHTSCPVGTCCPSLK.....QHCCPAGYTCNVKARTCEKQ 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	352	100.0	57	9 US-09-874-056-4	Sequence 4, Appli
2	347	98.6	58	9 US-09-874-056-5	Sequence 5, Appli
3	347	98.6	58	9 US-09-874-056-7	Sequence 7, Appli
4	347	98.6	589	9 US-09-824-647-2	Sequence 2, Appli
5	347	98.6	589	10 US-09-813-156-2	Sequence 2, Appli
6	347	98.6	589	10 US-09-824-807-2	Sequence 2, Appli
7	341	96.9	59	9 US-09-874-056-6	Sequence 6, Appli
8	338	96.0	621	10 US-09-925-301-1416	Sequence 1416, Ap
9	333	94.6	57	9 US-09-874-056-3	Sequence 3, Appli
10	325	92.3	593	9 US-09-824-647-17	Sequence 17, Appl
11	325	92.3	593	10 US-09-813-156-17	Sequence 17, Appl
12	325	92.3	593	10 US-09-824-807-17	Sequence 17, Appl
13	157	44.6	77	10 US-09-864-761-43653	Sequence 43653, A
14	93	26.4	53	9 US-09-874-056-8	Sequence 8, Appli
15	91.5	26.0	2476	9 US-10-184-644-585	Sequence 585, App
16	91.5	26.0	2476	9 US-10-184-634-585	Sequence 585, App
17	91.5	26.0	2738	9 US-10-184-644-575	Sequence 575, App
18	91.5	26.0	2738	9 US-10-184-634-575	Sequence 575, App
19	89	25.3	1781	9 US-10-123-155-419	Sequence 419, App

20	86	24.4	1606	9 US-10-123-155-65	Sequence 65, Appl
21	86	24.4	1606	9 US-10-123-155-257	Sequence 257, App
22	85.5	24.3	1685	9 US-10-123-155-347	Sequence 347, App
23	84	23.9	1863	9 US-10-184-644-547	Sequence 547, App
24	84	23.9	1863	9 US-10-184-634-547	Sequence 547, App
25	84	23.9	2292	9 US-10-184-644-493	Sequence 493, App
26	84	23.9	2292	9 US-10-184-634-493	Sequence 493, App
27	84	23.9	3162	9 US-10-123-155-111	Sequence 111, App
28	83.5	23.7	4277	9 US-10-184-644-439	Sequence 439, App
29	83.5	23.7	4277	9 US-10-184-634-439	Sequence 439, App
30	82.5	23.4	174	10 US-09-764-898-189	Sequence 189, App
31	82.5	23.4	500	9 US-10-241-476-2	Sequence 2, Appli
32	82.5	23.4	659	9 US-10-241-476-3	Sequence 3, Appli
33	82.5	23.4	685	9 US-10-028-072-88	Sequence 88, Appl
34	82.5	23.4	685	9 US-10-121-049-88	Sequence 88, Appl
35	82.5	23.4	685	9 US-10-123-904-88	Sequence 88, Appl
36	82.5	23.4	685	9 US-10-140-470-88	Sequence 88, Appl
37	82.5	23.4	685	9 US-10-175-746-88	Sequence 88, Appl
38	82.5	23.4	685	9 US-10-176-918-88	Sequence 88, Appl
39	82.5	23.4	685	9 US-10-176-921-88	Sequence 88, Appl
40	82.5	23.4	685	9 US-10-227-884-214	Sequence 214, App
41	82.5	23.4	685	9 US-10-137-865-88	Sequence 88, Appl
42	82.5	23.4	685	9 US-10-140-474-88	Sequence 88, Appl
43	82.5	23.4	685	9 US-10-142-431-88	Sequence 88, Appl
44	82.5	23.4	685	9 US-10-143-114-88	Sequence 88, Appl
45	82.5	23.4	685	9 US-10-230-163-214	Sequence 214, App

ALIGNMENTS

RESULT 1

US-09-874-056-4

; Sequence 4, Application US/09874056

; Publication No. US20020192704A1

; GENERAL INFORMATION:

; APPLICANT: OKANO, Akira

; APPLICANT: ETO, Yuzuru

; TITLE OF INVENTION: Insulin Receptor-Related Receptor Binding Protein and Utilizat

; TITLE OF INVENTION: Same

; FILE REFERENCE: 209427050

; CURRENT APPLICATION NUMBER: US/09/874,056

; CURRENT FILING DATE: 2001-06-06

; PRIOR FILING DATE: JP 2000-170912

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 57

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-874-056-4

Query Match 100.0%; Score 352; DB 9; Length 57;

Best Local Similarity 100.0%; Pred. No. 5.5e-26;

Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGTCCPSLKSGWACCQLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57

Db 1 IGCQHTSCPVGTCCPSLKSGWACCQLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57

RESULT 2

US-09-874-056-5

; Sequence 5, Application US/09874056

; Publication No. US20020192704A1

; GENERAL INFORMATION:

; APPLICANT: OKANO, Akira

; APPLICANT: ETO, Yuzuru

; TITLE OF INVENTION: Insulin Receptor-Related Receptor Binding Protein and Utilizat

; TITLE OF INVENTION: Same

```
; ORGANISM: Mouse epithelin/granulin
US-09-824-647-2

Query Match          98.6%; Score 347; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 9.9e-25;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   1  IGDQHTSCPVGTCPCPSLKGSLGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
      |||||||
Db    440 IGDQHTSCPVGTCPCPSLKGSLGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 5
US-09-813-156-2
; Sequence 2, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813,156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-813-156-2

Query Match          98.6%; Score 347; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 9.9e-25;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   1  IGDQHTSCPVGTCPCPSLKGSLGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
      |||||||
Db    440 IGDQHTSCPVGTCPCPSLKGSLGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 6
US-09-824-807-2
; Sequence 2, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-807-2

Query Match          98.6%; Score 347; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 9.9e-25;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   1  IGDQHTSCPVGTCPCPSLKGSLGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
      |||||||
Db    440 IGDQHTSCPVGTCPCPSLKGSLGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495
```

```
; TITLE OF INVENTION: Same
; FILE REFERENCE: 209427US0
; CURRENT APPLICATION NUMBER: US/09/874,056
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: JP 2000-170912
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-874-056-3

Query Match          94.6%; Score 333; DB 9; Length 57;
Best Local Similarity 96.4%; Pred.No.3e+24; 2; Indels 0; Gaps 0;
Matches 54; Conservative 0; Mismatches 2;

QY 1 IGCDDHTSCPVGQTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
      |||
Db 2 IGCDDHTSCPVGQTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCGK 57
      |||

RESULT 10
US-09-824-647-17
; Sequence 17, Application US/09824647
; Publication No. US20020183270A1
; GENERAL INFORMATION:
; APPLICANT: Serreto, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,647
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 CDNA
US-09-824-647-17

Query Match          92.3%; Score 325; DB 9; Length 593;
Best Local Similarity 91.2%; Pred.No.1e+22;
Matches 52; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IGCDDHTSCPVGQTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
      |||
Db 442 IGCDDHTSCPVGGTCCPSQGSGWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498
      |||

RESULT 11
US-09-813-156-17
; Sequence 17, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serreto, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813,156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43653
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003043.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.65
; OTHER INFORMATION: EST_HUMAN HIT: BF344549.1, EVALUE 8.00e-46
; OTHER INFORMATION: SWISSPROT HIT: P28799, EVALUE 8.00e-47
US-09-864-761-43653

Query Match 44.6%; Score 157; DB 10; Length 77;
Best Local Similarity 92.6%; Pred. No. 5e-08;
Matches 25; Conservative 2; Mismatches 0; Indels

QY 31 AVCCEDRHQCCPAGYTCNVKARTCEKQ 57
      |||||
Db 1 AVCCEDRHQCCPAGYTCNVKARSEKE 27

RESULT 14
US-09-874-056-8
; Sequence 8, Application US/09874056
; Publication No. US20020192704A1
; GENERAL INFORMATION:
; APPLICANT: OKANO, Akira
; APPLICANT: ETO, Yuzuru
; APPLICANT: IZUMI, Tetsuro
; TITLE OF INVENTION: Insulin Receptor-Related Receptor Binding
; FILE OF INVENTION: Same
; FILE REFERENCE: 209427U50
; CURRENT APPLICATION NUMBER: US/09/874,056
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: JP 2000-170912
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
; FEATURE:
; NAME/KEY: MISC_FEATURE

```

```
LOCATION: (4)..(4)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (5)..(5)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (7)..(7)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (10)..(10)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (11)..(11)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (12)..(12)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (13)..(13)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (16)..(16)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (17)..(17)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (18)..(18)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (19)..(19)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (20)..(20)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (21)..(21)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (22)..(22)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (23)..(23)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (26)..(26)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (27)..(27)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (28)..(28)
```

```
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (29)..(29)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (30)..(30)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (31)..(31)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (34)..(34)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (36)..(36)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (37)..(37)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (42)..(42)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (43)..(43)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (44)..(44)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (45)..(45)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (47)..(47)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (48)..(48)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (49)..(49)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (50)..(50)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (51)..(51)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (52)..(52)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (9)..(9)
OTHER INFORMATION: X IS ANY AMINO ACID OR NONEXISTENT
```

US-09-874-056-8

Query Match 26.4%; Score 93; DB 9; Length 53;
Best Local Similarity 30.4%; Pred. No. 0.027;

9 CPVGQTCCPSLKGSWACCCQLPHAVCCEDRQHCCPAGYTCNVKARTC 54
QY
8 CXXXXXCCXXXXXXXXXXCCXXXXXXXXCCDXHCCPXXXXXXXXXXXXC 53
Db

US-10-184-644-585

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

1 / TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
2 /
3 / TITLE OF INVENTION: ACIDS ENCODING THE SAME
4 /

; CURRENT APPLICATION NUMBER: US/10/184,644

; Prior Application removed - See

; SEQ ID NO 585

TYPE: DNA

ORGANISM: Homo sapiens
US-10-184-644-585

Query Match 26 0% score 91 5: DB 9: Tenath 2476:

BEST LOCAL SIMILARITY MATCHES 20: Conserv

3 CPOHTCOPV---COTCOPCT.KGSWACCOI.PHAYCCEPPOHCCPACYTCHVKABTC 5A

[illegible]

Search completed: April 21, 2003. 14:24:33

Job time : 16 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:20:05 ; Search time 28 seconds
(without alignments)
59.897 Million cell updates/sec

Title: US-09-874-056-4
Perfect score: 352
Sequence: 1 IGCDOHTSCPVGTCCPSLK.....QHCCPAGYTCNVKARTCEK 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	347	98.6	589	1 US-07-668-648-2	Sequence 2, Appli
2	347	98.6	589	1 US-07-668-648-6	Sequence 6, Appli
3	347	98.6	589	2 US-08-429-998-2	Sequence 2, Appli
4	347	98.6	589	2 US-08-429-998-6	Sequence 6, Appli
5	347	98.6	589	2 US-08-431-333-2	Sequence 2, Appli
6	347	98.6	589	2 US-08-431-333-6	Sequence 6, Appli
7	347	98.6	589	4 US-08-991-862-2	Sequence 2, Appli
8	347	98.6	589	5 PCT-US91-02321-2	Sequence 2, Appli
9	347	98.6	589	5 PCT-US91-02321-6	Sequence 6, Appli
10	338	96.0	593	1 US-07-668-648-4	Sequence 4, Appli
11	338	96.0	593	2 US-08-429-998-4	Sequence 4, Appli
12	338	96.0	593	2 US-08-431-333-4	Sequence 4, Appli
13	338	96.0	593	5 PCT-US91-02321-4	Sequence 4, Appli
14	325	92.3	593	4 US-08-991-862-17	Sequence 17, Appli
15	181	51.4	179	1 US-07-668-648-8	Sequence 8, Appli
16	181	51.4	179	2 US-08-429-998-8	Sequence 8, Appli
17	181	51.4	179	2 US-08-431-333-8	Sequence 8, Appli
18	181	51.4	179	5 PCT-US91-02321-8	Sequence 8, Appli
19	153	43.5	113	1 US-07-668-648-10	Sequence 10, Appli
20	153	43.5	113	2 US-08-429-998-10	Sequence 10, Appli
21	153	43.5	113	2 US-08-431-333-10	Sequence 10, Appli
22	153	43.5	113	5 PCT-US91-02321-10	Sequence 10, Appli
23	150	42.6	500	4 US-09-325-932A-149	Sequence 149, App
24	139	39.5	443	4 US-08-821-99A-65	Sequence 65, Appli
25	136	38.6	457	3 US-09-120-365-72	Sequence 72, Appli
26	136	38.6	457	4 US-09-515-039-72	Sequence 72, Appli
27	127	36.1	442	4 US-08-821-99A-66	Sequence 66, Appli

28	84.5	24.0	635	1	US-07-907-190-1	Sequence 1, Appli
29	84.5	24.0	635	1	US-07-985-691-2	Sequence 2, Appli
30	84.5	24.0	635	1	US-08-436-804-2	Sequence 2, Appli
31	84.5	24.0	635	1	US-08-267-387-2	Sequence 2, Appli
32	84.5	24.0	676	1	US-08-282-141-4	Sequence 4, Appli
33	84.5	24.0	676	1	US-08-435-434-3	Sequence 3, Appli
34	84.5	24.0	676	1	US-08-435-436-3	Sequence 3, Appli
35	84.5	24.0	676	2	US-08-438-863-3	Sequence 3, Appli
36	84.5	24.0	676	2	US-08-438-864-3	Sequence 3, Appli
37	84.5	24.0	676	2	US-08-438-862-3	Sequence 3, Appli
38	84.5	24.0	676	4	US-08-628-747-3	Sequence 3, Appli
39	84.5	24.0	676	4	US-08-402-253-3	Sequence 3, Appli
40	84.5	24.0	676	4	US-08-443-866B-3	Sequence 3, Appli
41	83.5	23.7	713	3	US-08-872-855-5	Sequence 5, Appli
42	82.5	23.4	685	3	US-08-872-855-2	Sequence 2, Appli
43	82.5	23.4	720	3	US-08-872-855-4	Sequence 4, Appli
44	82.5	23.4	722	4	US-08-981-392-12	Sequence 12, Appli
45	82	23.3	578	4	US-08-981-392-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-07-668-648-2
; Sequence 2, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-668-648-2

Query Match 98.6%; Score 347; DB 1; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGCDOHTSCPVGTCCPSLKSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
DB 440 IGCDOHTSCPVGTCCPSLKSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

3


```
Query Match          98.6%; Score 347; DB 2; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 5
US-08-431-333-2
; Sequence 2, Application US/08431333
; Patent No. 5965723
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 27-APR-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,333
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/668,648
; FILING DATE: 13-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-431-333-2

Query Match          98.6%; Score 347; DB 2; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 6
US-08-431-333-6
; Sequence 6, Application US/08431333
; Patent No. 5965723
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,333
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/668,648
; FILING DATE: 13-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-431-333-6

Query Match          98.6%; Score 347; DB 2; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 1 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 7
US-08-991-862-2
; Sequence 2, Application US/08991862
; Patent No. 6309826
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/08/991,862
; CURRENT FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 08/863,862
; EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-08-991-862-2

Query Match          98.6%; Score 347; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCDOHTSCPVGOTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 8
```

PCT-US91-02321-2
; Sequence 2, Application PC/TUS9102321
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02321
; FILING DATE: 19910403
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: ON0071A-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)728-4800
; TELEFAX: (206)448-4775
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-02321-2

Query Match 98.6%; Score 347; DB 5; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPCSLKSGWACCQLPHAVCCEDRHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCQHTSCPVGQTCPCSLKSGWACCQLPHAVCCEDRHCCPAGYTCNVKARTCEK 495

RESULT 9
PCT-US91-02321-6
; Sequence 6, Application PC/TUS9102321
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02321
; FILING DATE: 19910403
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: ON0071A-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)728-4800
; TELEFAX: (206)448-4775
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-02321-6

Query Match 98.6%; Score 347; DB 5; Length 589;

Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPCSLKSGWACCQLPHAVCCEDRHCCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCQHTSCPVGQTCPCSLKSGWACCQLPHAVCCEDRHCCPAGYTCNVKARTCEK 495

RESULT 10

us-07-668-648-4
; Sequence 4, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-668-648-4

Query Match 96.0%; Score 338; DB 1; Length 593;

Best Local Similarity 94.7%; Pred. No. 1.1e-26;
Matches 54; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPCSLKSGWACCQLPHAVCCEDRHCCPAGYTCNVKARTCEK 57
|||||
Db 442 IGCQHTSCPVGQTCPCSLKSGWACCQLPHAVCCEDRHCCPAGYTCNVKARTCEK 498

RESULT 11

US-08-429-998-4
; Sequence 4, Application US/0842998
; Patent No. 5885961
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: PLOWMAN, Gregory D.
; TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,998

; FILING DATE: 27-APR-1995

; CLASSIFICATION: 514

; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/668,648

; FILING DATE: 13-MAR-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 5624-161-999

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090

; TELEFAX: (212) 869-9741

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 593 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-429-998-4

Query Match 96.0%; Score 338; DB 2; Length 593;
Best Local Similarity 94.7%; Pred. No. 1.le-26;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
|||||
Db 442 IGCQHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498

RESULT 12

US-08-431-333-4
; Sequence 4, Application US/08431333
; Patent No. 5965723
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: PLOWMAN, Gregory D.
; TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,333

; FILING DATE: 27-APR-1995

; CLASSIFICATION: 536

; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/668,648

; FILING DATE: 13-MAR-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 5624-161-999

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090

; TELEFAX: (212) 869-9741

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 593 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-431-333-4

Query Match 96.0%; Score 338; DB 2; Length 593;

Best Local Similarity 94.7%; Pred. No. 1.le-26;

Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IGCQHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
|||||

Db 442 IGCQHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498

RESULT 13

PCT-US91-02321-4

; Sequence 4, Application PC/TUS9102321

; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed

; APPLICANT: PLOWMAN, Gregory D.

; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH

; TITLE OF INVENTION: MODULATING PROTEINS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company

; STREET: 3005 First Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98121

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02321

; FILING DATE: 19910403

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.

; REGISTRATION NUMBER: 32,928

; REFERENCE/DOCKET NUMBER: ON0071A-PC

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)728-4800

; TELEFAX: (206)448-4775

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 593 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US91-02321-4

Query Match 96.0%; Score 338; DB 5; Length 593;
Best Local Similarity 94.7%; Pred. No. 1.1e-26;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
|||||
DB 442 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498
|||||

RESULT 14

US-08-991-862-17
; Sequence 17, Application US/08991862
; Patent No. 6309826
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/08/991,862
; CURRENT FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 08/863,862
; EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 CDNA
; US-08-991-862-17

Query Match 92.3%; Score 325; DB 4; Length 593;
Best Local Similarity 91.2%; Pred. No. 2.3e-25;
Matches 52; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
|||||
DB 442 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498
|||||

RESULT 15

US-07-668-648-8
; Sequence 8, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; - APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 8;

; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-668-648-8

Query Match 51.4%; Score 181; DB 1; Length 179;
Best Local Similarity 46.4%; Pred. No. 2e-11;
Matches 26; Conservative 7; Mismatches 23; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGQTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
: || ||| ||| : || ||| ||| ||| ||| : |||
DB 54 VKCDMEVSCPDYTCRLQSGAWGCCPFVQAVCCEDHVHCCPSGFRGDTKGVCEQ 109

Search completed: April 21, 2003, 14:24:11
Job time : 29 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:11:40 ; Search time 35 Seconds
(without alignments)
217.008 Million cell updates/sec

Title: US-09-874-056-4

Perfect score: 352

Sequence: 1 IGCQHTSCVPVGTCCPSLK.....QHCCPAGYTCNVKARTCEKQ 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	347	98.6	57	ABB07323	Insulin receptor-r
2	347	98.6	57	ABB07324	Insulin receptor-r
3	347	98.6	58	ABB07325	Insulin receptor-r
4	347	98.6	58	ABB07327	Insulin receptor-r
5	347	98.6	59	ABB07326	Insulin receptor-r
6	347	98.6	589	12 AAR14325	Rat epithelin prec
7	347	98.6	589	12 AAR14327	Mouse epithelin pr
8	347	98.6	589	23 AAE20520	Mouse granulin/epi
9	345	98.0	589	20 AAW85474	Mouse GP88 autocri
10	338	96.0	413	22 AAB94550	Human protein sequ

11	338	96.0	593	12 AAR14326	Human epithelin pr
12	338	96.0	621	21 AAB43971	Human cancer assoc
13	325	92.3	593	14 AAR48673	Granulin sequence.
14	325	92.3	593	20 AAW85475	Human GP88 autocri
15	325	92.3	593	23 AAE20521	Human granulin/epi
16	318	90.3	55	14 AAR41882	Granulin D. Homo
17	190.5	54.1	55	14 AAR41879	Granulin A. Homo
18	182	51.7	54	14 AAR41881	Granulin C. Homo
19	181	51.4	179	12 AAR15426	Bovine epithelin p
20	178	50.6	51	14 AAR41883	Rat granulin. Rat
21	162.5	46.2	57	14 AAR41886	Granulin F. Homo
22	157	44.6	77	22 AAM59508	Human brain expres
23	157	44.6	469	20 AAW89560	Triticum sp. cyste
24	156	44.3	472	20 AAW89559	Triticum sp. cyste
25	153	43.5	113	12 AAR15427	Chicken epithelin
26	152	43.2	56	14 AAR41885	Granulin E. Homo
27	150	42.6	500	22 AAB65766	Cysteine protease
28	145	41.2	56	14 AAR41880	Granulin B. Homo
29	143	40.6	256	21 AAG07186	Arabidopsis thalia
30	143	40.6	325	21 AAG07185	Arabidopsis thalia
31	143	40.6	344	21 AAG39794	Arabidopsis thalia
32	143	40.6	411	21 AAG39793	Arabidopsis thalia
33	143	40.6	452	21 AAG39792	Arabidopsis thalia
34	141	40.1	266	21 AAG35415	Zea mays protein f
35	141	40.1	419	21 AAG35414	Zea mays protein f
36	141	40.1	424	21 AAG35413	Zea mays protein f
37	139	39.5	443	18 AAW27441	Oil seed rape cyst
38	138	39.2	461	20 AAW89558	Triticum sp. cyste
39	134	38.1	464	18 AAW19542	Soybean thiol prot
40	134	38.1	464	23 ABB81808	Soybean D3-beta.
41	130	36.9	132	21 AAG09931	Arabidopsis thalia
42	130	36.9	155	21 AAG09930	Arabidopsis thalia
43	130	36.9	463	18 AAW19541	Soybean thiol prot
44	127	36.1	442	18 AAW27442	Oil seed rape cyst
45	122	34.7	54	14 AAR41887	Granulin G. Homo

ALIGNMENTS

RESULT 1

ABB07323
ID ABB07323 standard; protein; 57 AA.

XX ABB07323;

XX ABB07323;

DT 09-APR-2002 (first entry)

XX Insulin receptor-related receptor binding protein fragment.

DE IRR: antidiabetic; vulnery; neuroprotective; nephrotropic; cytostatic;

KW insulin receptor-related receptor; mouse; epithelin; granulin precursor.

XX Mus musculus.

OS Mus musculus.

XX EP1164144-A1.

PD 19-DEC-2001.

XX 07-JUN-2001; 2001EP-0112975.

XX 07-JUN-2000; 2000JP-0170912.

XX (AJIN) AJINMOTO CO INC.

XX Okano A, Eto Y, Izumi T;

XX WPI; 2002-124195/17.

XX New insulin receptor-related receptor (IRR) binding protein for

XX identifying an agonist and antagonist, or for use in the treatment of

XX diabetes -

CC living body. IRR is expressed in specific localized tissues and cells
 CC such as in the kidney, nerve and stomach, and is thought to be involved
 CC in growth/differentiation of these cells. (I), or an agonist, could
 CC therefore be used as a therapeutic agent for congenital or acquired
 CC neuropathy, renal disorder, gastrointestinal injury due to medication,
 CC autoimmunity or inflammation. Also an antagonist of (I) could be
 CC therapeutic for diseases associated with hyperplasia or hyperactivity
 CC of the tissues or cells mentioned, e.g., proliferative nephritis. The
 CC present sequence represents a IRR binding protein fragment that is
 CC homologous to a rat epithelin/granulin precursor.

SQ Sequence 58 AA;
 Query Match 98.6%; Score 347; DB 23; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGTCTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
 DB 2 IGCQHTSCPVGTCTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 57

RESULT 4
 ABB07327
 ID ABB07327 standard; protein; 58 AA.

XX ABB07327;
 XX 09-APR-2002 (first entry)
 XX Insulin receptor-related receptor binding protein fragment.
 DE IRR; antidiabetic; vulnery; neuroprotective; nephrotropic; cytostatic;
 KW insulin receptor-related receptor; mouse; epithelin; granulin precursor.
 XX Mus musculus.

XX EP1164144-A1.
 XX 19-DEC-2001.
 XX 07-JUN-2001; 2001EP-0112975.
 XX 07-JUN-2000; 2000JP-0170912.
 XX (AJIN) AJINMOTO CO INC.
 XX Okano A, Eto Y, Izumi T;
 XX WPI; 2002-124195/17.

XX New insulin receptor-related receptor (IRR) binding protein for
 PT identifying an agonist and antagonist, or for use in the treatment of
 PT diabetes -

PS Claim 2; Page 10; 21pp; English.

XX The invention relates to a protein (I), or variant, from Mus musculus
 CC which has binding activity to an insulin receptor-related receptor (IRR).
 CC (I) has a molecular weight of 6135, 6206, 6250 or 6321 as measured by
 CC mass spectrometry using the Fourier transformation ion cyclotron method.
 CC Compositions containing (I) can regulate growth/differentiation in cells
 CC related in diabetes, neuropathy, renal disorder or gastrointestinal
 CC injury and could therefore be used as a therapeutic agent in such cases.
 CC IRR is highly expressed in pancreatic beta cells so (I) could be used
 CC as a therapeutic agent in the treatment of diabetes via regulation of
 CC growth/differentiation of pancreatic beta cells when administered to a
 CC living body. IRR is expressed in specific localized tissues and cells
 CC such as in the kidney, nerve and stomach, and is thought to be involved
 CC in growth/differentiation of these cells. (I), or an agonist, could
 CC therefore be used as a therapeutic agent for congenital or acquired
 CC neuropathy, renal disorder, gastrointestinal injury due to medication,
 CC autoimmunity or inflammation. Also an antagonist of (I) could be

CC therapeutic for diseases associated with hyperplasia or hyperactivity
 CC of the tissues or cells mentioned, e.g., proliferative nephritis. The
 CC present sequence represents a IRR binding protein fragment that is
 CC homologous to a rat epithelin/granulin precursor.

SQ Sequence 58 AA;

Query Match 98.6%; Score 347; DB 23; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGTCTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
 DB 1 IGCQHTSCPVGTCTCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56

RESULT 5
 ABB07326
 ID ABB07326 standard; protein; 59 AA.

XX ABB07326;
 XX 09-APR-2002 (first entry)
 XX Insulin receptor-related receptor binding protein fragment.
 DE IRR; antidiabetic; vulnery; neuroprotective; nephrotropic; cytostatic;
 KW insulin receptor-related receptor; mouse; epithelin; granulin precursor.
 XX Mus musculus.

XX EP1164144-A1.
 XX 19-DEC-2001.
 XX 07-JUN-2001; 2001EP-0112975.
 XX 07-JUN-2000; 2000JP-0170912.
 XX (AJIN) AJINMOTO CO INC.
 XX Okano A, Eto Y, Izumi T;
 XX WPI; 2002-124195/17.

XX New insulin receptor-related receptor (IRR) binding protein for
 PT identifying an agonist and antagonist, or for use in the treatment of
 PT diabetes -

PS Claim 2; Page 10; 21pp; English.

XX The invention relates to a protein (I), or variant, from Mus musculus
 CC which has binding activity to an insulin receptor-related receptor (IRR).
 CC (I) has a molecular weight of 6135, 6206, 6250 or 6321 as measured by
 CC mass spectrometry using the Fourier transformation ion cyclotron method.
 CC Compositions containing (I) can regulate growth/differentiation in cells
 CC related in diabetes, neuropathy, renal disorder or gastrointestinal
 CC injury and could therefore be used as a therapeutic agent in such cases.
 CC IRR is highly expressed in pancreatic beta cells so (I) could be used
 CC as a therapeutic agent in the treatment of diabetes via regulation of
 CC growth/differentiation of pancreatic beta cells when administered to a
 CC living body. IRR is expressed in specific localized tissues and cells
 CC such as in the kidney, nerve and stomach, and is thought to be involved
 CC in growth/differentiation of these cells. (I), or an agonist, could
 CC therefore be used as a therapeutic agent for congenital or acquired
 CC neuropathy, renal disorder, gastrointestinal injury due to medication,
 CC autoimmunity or inflammation. Also an antagonist of (I) could be
 CC therapeutic for diseases associated with hyperplasia or hyperactivity
 CC of the tissues or cells mentioned, e.g., proliferative nephritis. The
 CC present sequence represents a IRR binding protein fragment that is
 CC homologous to a rat epithelin/granulin precursor.

XX Sequence 59 AA;

Query Match 98.6%; Score 347; DB 23; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.6e-25;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGQTCPCSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||
Db 2 IGCQHTSCPVGQTCPCSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 57

RESULT 6
AAR14325
ID AAR14325 standard; Protein; 589 AA.

AC AAR14325;

XX 17-JAN-1992 (first entry)

XX Rat epithelin precursor.

XX ET; growth regulation; inhibition; stimulation.

XX Rattus rattus.

Key Location/Qualifiers

FT Protein 1..589
FT /label= precursor
FT /note= "claim 11, page 54"
FT Protein 280..335
FT /label= EP-1
FT /note= "claim 12, page 54"
FT Protein 205..261
FT /label= EP-2
FT /note= "claim 13, page 54"
FT Peptide 59..114
FT /label= EP
FT /note= "claim 14, page 54"
FT Peptide 123..179
FT /label= EP
FT /note= "claim 15, page 54"
FT Peptide 362..416
FT /label= EP
FT /note= "claim 16, page 54"
FT Peptide 440..495
FT /label= EP
FT /note= "claim 17, page 54"
FT Peptide 515..570
FT /label= EP
FT /note= "claim 18, page 55"

PN WO9115510-A.

XX 17-OCT-1991.

XX 03-APR-1991; 91WO-US02321.

XX 13-MAR-1991; 91US-0083796.

PR 03-APR-1990; 90US-0504508.

XX (BRIM) BRISTOL-MYERS SQUIB.

XX Shoyab M, Plowman GD;

XX WPI; 1991-325168/44.

DR N-PSDB; AAQ14338.

XX New cysteine-rich growth modulating proteins, epithelins - useful
PT as inhibitors of neoplastic cell growth and to promote wound
PT healing and treat psoriasis

XX Disclosure; Fig 18; 97pp; English.

XX ET-1 is a bifunctional growth regulator, capable of stimulating

CC the growth of some cell types while inhibiting the growth of others.
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
CC bioactivity. In contrast, however, ET-2 is apparently not capable of
CC eliciting the growth stimulatory activity characteristic of ET-1 and,
CC in fact, antagonizes this ET-1 activity.
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.

XX Sequence 589 AA;

Query Match 98.6%; Score 347; DB 12; Length 589;

Best Local Similarity 100.0%; Pred. No. 1.8e-24;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCQHTSCPVGQTCPCSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56
|||||

Db 440 IGCQHTSCPVGQTCPCSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 7

AAR14327

ID AAR14327 standard; Protein; 589 AA.

XX AAR14327;

XX 17-JAN-1992 (first entry)

XX Mouse epithelin precursor.

XX ET; growth regulation; inhibition; stimulation.

XX Mus musculus.

Key Location/Qualifiers

FT Protein 1..589
FT /label= precursor
FT /note= "claim 21, page 55"
FT Protein 280..335
FT /label= EP-1
FT /note= "claim 22, page 55"
FT Protein 205..261
FT /label= EP-2
FT /note= "claim 23, page 55"
FT Peptide 59..114
FT /label= EP
FT /note= "claim 24, page 55"
FT Peptide 123..179
FT /label= EP
FT /note= "claim 25, page 55"
FT Peptide 362..416
FT /label= EP
FT /note= "claim 26, page 56"
FT Peptide 440..495
FT /label= EP
FT /note= "claim 27, page 56"
FT Peptide 515..570
FT /label= EP
FT /note= "claim 28, page 56"

PN WO9115510-A.

XX 17-OCT-1991.

XX 03-APR-1991; 91WO-US02321.

XX 13-MAR-1991; 91US-0083796.

PR 03-APR-1990; 90US-0504508.

XX (BRIM) BRISTOL-MYERS SQUIB.

XX Shoyab M, Plowman GD;

XX WPI; 1991-325168/44.

DR N-PSDB; AAQ14340.

XX New cysteine-rich growth modulating proteins, epithelins - useful
PT as inhibitors of neoplastic cell growth and to promote wound
PT healing and treat psoriasis
XX
XX Disclosure; Fig 23; 97pp; English.
CC
CC ET-1 is a bifunctional growth regulator, capable of stimulating
CC the growth of some cell types while inhibiting the growth of others.
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
CC bioactivity. In contrast, however, ET-2 is apparently not capable of
CC eliciting the growth stimulatory activity characteristic of ET-1 and,
CC in fact, antagonises this ET-1 activity.
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.
XX
SQ Sequence 589 AA;
Query Match 98.6%; Score 347; DB 12; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.8e-24;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGCDHTSCPVGQTCCPSLKGSWACQLPHAVCCEDRHCPCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCDHTSCPVGQTCCPSLKGSWACQLPHAVCCEDRHCPCPAGYTCNVKARTCEK 495
RESULT 8
AAE20520
ID AAE20520 standard; Protein; 589 AA.
XX
AC AAE20520;
XX
DT 01-JUL-2002 (first entry)
XX
XX Mouse granulin/epithelin precursor (GP88) protein.
XX
XX Mouse; granulin precursor; GP88; cytostatic; tumourigenicity; tamoxifen;
XX antineoplastic; antioestrogen therapy; skin cancer.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 335
FT /label= Unknown
FT /note= "Encoded by ATG"
FT Region 344..362
FT /note= "Regions used as antigens to raise anti-GP88
FT antibodies"
FT Region 562..575
FT /note= "Regions used as antigens to raise anti-GP88
FT antibodies"
FT Misc-difference 586
FT /note= "Encoded by ACA"
XX
XX US2002025543-A1.
PN
XX
XX 28-FEB-2002.
PD
XX
XX 15-JUN-2001; 2001US-0880842.
PF
XX
XX 23-MAY-1997; 97US-0863079.
PR
XX 08-DEC-1999; 99US-0456886.
PR
XX
XX (SERR/) SERRERO G.
PA
XX
XX Serrero G;
PI
XX
XX WPI; 2002-267529/31.
DR
XX N-PSDB; AAD32849.
DR
XX
XX Diagnosing tumorigenicity in a human, comprising obtaining a cell
PT sample, detecting GP88 in the cells, and determining the number of GP88
PT positive cells in the sample

XX Disclosure; Fig 8; 50pp; English.
PS
XX
CC The invention relates to a method for diagnosing tumourigenicity in a
CC human. The method comprises obtaining a biological sample containing
CC cells from the patient, detecting GP88 in the cells of the sample, and
CC determining the number of GP88 positive cells in the sample, and
CC determining the ratio of GP88 positive cells to the total number of cells
CC in the sample. The invention also relates to a method for determining
CC if a human patient is resistant to the antineoplastic effects of
CC antioestrogen therapy. The method is useful for diagnosing
CC tumourigenicity in a sample, such as blood, serum, plasma, urine, nipple
CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,
CC colon, or skin cancer. The method can be used to treat or prevent
CC re-occurrence of cancer in a patient, by administering tamoxifen if the
CC sample contains less than 10 % GP88, or less than 5 % GP88 positive
CC cells. The present sequence is mouse granulin/epithelin precursor (GP88)
CC protein.
XX
SQ Sequence 589 AA;
Query Match 98.6%; Score 347; DB 23; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.8e-24;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGCDHTSCPVGQTCCPSLKGSWACQLPHAVCCEDRHCPCPAGYTCNVKARTCEK 56
|||||
Db 440 IGCDHTSCPVGQTCCPSLKGSWACQLPHAVCCEDRHCPCPAGYTCNVKARTCEK 495
RESULT 9
AAW85474
ID AAW85474 standard; Protein; 589 AA.
XX
AC AAW85474;
XX
DT 15-MAR-1999 (first entry)
XX
XX Mouse GP88 autocrine growth factor.
XX
XX GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;
KW cancer; viral infection; antagonist; therapy; diagnosis.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 8
FT /note= "encoded by CTG"
FT Misc-difference 54
FT /note= "encoded by AGC"
FT Misc-difference 377
FT /note= "encoded by TGA"
FT Peptide 208..219
FT /note= "p12r peptide used to raise antibody"
FT Peptide 344..362
FT /note= "K19T peptide, used to raise antibody"
FT Peptide 562..575
FT /note= "S14R peptide, used to raise antibody"
XX
XX WO9852607-A1.
PN
XX
XX 26-NOV-1998.
PD
XX
XX 22-MAY-1998; 98WO-US10555.
PF
XX
XX 16-DEC-1997; 97US-0991862.
PR
XX 23-MAY-1997; 97US-0863079.
PR
XX
XX (SERR/) SERRERO G.
PA
XX
XX Serrero G;
PI
XX
XX WPI; 1999-045276/04.
DR

DR N-PSDB; AAV82824.

XX Composition containing antagonist of growth factor gp88 - useful for

PT treating cancer and viral diseases and also for diagnosing disease

PT from altered GP88 expression

XX

XX Example 5; Fig 8A-D; 86pp; English.

XX This is the amino acid sequence of murine GP88, an 88 kDa

CC glycoprotein autocrine growth factor and epithelin/granulin

CC precursor that is expressed in a tightly regulated manner in normal

CC cells, is overexpressed and unregulated in highly tumorigenic cells

CC derived from normal cells, and which acts as a stringently required

CC growth stimulator for the tumorigenic cells. Inhibition of GP88

CC expression or action in the tumorigenic cells results in an

CC inhibition of the tumorigenic properties of the overproducing

CC cells. Murine GP88 cDNA (see AAV82824) was obtained from the highly

CC tumorigenic PC cell line. Antagonists to GP88 are used to treat

CC diseases associated with increased expression of GP88, particularly

CC cancer but also viral infections. Fragments of GP88 are used to

CC raise specific antibodies (used as antagonists, as diagnostic

CC reagents and for delivering toxins or other compounds to GP88-

CC expressing cells) and to screen for antibodies. Methods are

CC provided for diagnosing disease, or determining susceptibility to

CC disease, resulting from altered GP88 activity.

XX

XX Sequence 589 AA;

SQ

Query Match 98.0%; Score 345; DB 20; Length 589;

Best Local Similarity 98.2%; Pred. No. 2.7e-24;

Matches 55; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGCDOHTSCPVGOTCCPSLKGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 56

DB 440 LGCDQHTSCPVGOTCCPSLKGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 495

RESULT 10

AA94550

ID AAB94550 standard; Protein; 413 AA.

XX

AC AAB94550;

XX

XX 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15310.

XX

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-0116126.

XX

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

XX (HELI-) HELIX RES INST.

XX

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

XX WPI; 2001-318749/34.

XX

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX

PS Claim 8; SEQ ID 15310; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

SQ Sequence 413 AA;

Query Match 96.0%; Score 338; DB 22; Length 413;

Best Local Similarity 94.7%; Pred. No. 9e-24;

Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IGCDOHTSCPVGOTCCPSLKGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 57

DB 262 IGCDOHTSCPVGOTCCPSLKGWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEK 318

RESULT 11

AA14326

ID AAR14326 standard; Protein; 593 AA.

XX

AC AAR14326;

XX

XX 17-JAN-1992 (first entry)

DE Human epithelin precursor.

XX

XX ET; growth regulation; inhibition; stimulation.

XX

OS Homo sapiens.

XX

PH Key

FT Location/Qualifiers

FT 1..593

FT /label= precursor

FT /note= "claim 1, page 53"

FT 282..337

FT Protein

FT /label= EP-1

FT /note= "claim 2, page 53"

FT 206..262

FT Protein

FT /label= EP-2

FT /note= "claim 3, page 53"

FT 59..114

FT Peptide

FT /label= EP

FT /note= "claim 4, page 53"

FT 124..180

FT Peptide

FT /label= EP

FT /note= "claim 5, page 53"

FT 364..418

FT Peptide

FT /label= EP

FT /note= "claim 6, page 53"

FT 442..497

FT Peptide

FT /label= EP

FT Peptide /note= "claim 7, page 53"
FT 519..574
FT /label= EP
FT /note= "claim 8, page 55"
XX
XX
XX WO9115510-A.
XX
XX 17-OCT-1991.
XX
XX 03-APR-1991; 91WO-US02321.
XX
XX 13-MAR-1991; 91US-0083796.
XX 03-APR-1990; 90US-0504508.
XX
XX (BRIM) BRISTOL-MYERS SQUIB.
XX
XX Shoyab M, Plowman GD;
XX WPI; 1991-325168/44.
XX N-PSDB; AAQ14339.
XX
XX New cysteine-rich growth modulating proteins, epithelins - useful
XX as inhibitors of neoplastic cell growth and to promote wound
XX healing and treat psoriasis
XX
XX Disclosure; Fig 22; 97pp; English.
XX
XX ET-1 is a bifunctional growth regulator, capable of stimulating
XX the growth of some cell types while inhibiting the growth of others.
XX ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
XX bioactivity. In contrast, however, ET-2 is apparently not capable of
XX eliciting the growth stimulatory activity characteristic of ET-1 and,
XX in fact, antagonises this ET-1 activity.
XX See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.
XX
XX Sequence 593 AA;
XX
Query Match 96.0%; Score 338; DB 12; Length 593;
Best Local Similarity 94.7%; Pred. No. 1.2e-23;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 IGCDDHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
Db 442 IGCDDHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498
RESULT 12
AAR43971
ID AAR43971 standard; Protein; 621 AA.
XX
XX AAR43971;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated protein sequence SEQ ID NO:1416.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
XX antinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening.
XX
XX Homo sapiens.
OS
XX WO200055350-A1.
XX
XX 21-SEP-2000.
XX

PF 08-MAR-2000; 2000WO-US05882.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX N-PSDB; AAC78180.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer -
XX
XX Claim 11; Page 2094-2096; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given
XX in AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnery; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
XX nootropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention.
XX
XX Sequence 621 AA;
XX
Query Match 96.0%; Score 338; DB 21; Length 621;
Best Local Similarity 94.7%; Pred. No. 1.3e-23;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 IGCDDHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
Db 470 IGCDDHTSCPVGQTCPCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 526
RESULT 13
AAR48673
ID AAR48673 standard; Protein; 593 AA.
XX
XX AAR48673;
XX
XX 22-APR-1994 (first entry)
XX
XX Granulin sequence.
XX
XX Granulin; keratinocytes; wound healing; inhibition; peptide;
XX granulocytes; leucocytes.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Key Misc-difference 452 /note= "valine encoded by ATG."
XX FT Misc-difference 539 /note= "Glycine encoded by CAG."
XX
XX WO9315195-A.
XX
XX 05-AUG-1993.
XX

XX	CC	This is the amino acid sequence of human GP88, an 88 kDa									
XX	CC	glycoprotein autocrine growth factor and epithelin/granulin									
XX	CC	precursor that is expressed in a tightly regulated manner in normal									
XX	CC	cells, is overexpressed and unregulated in highly tumorigenic cells									
XX	CC	derived from normal cells, and which acts as a stringently required									
XX	CC	growth stimulator for the tumorigenic cells. Inhibition of GP88									
XX	CC	expression or action in the tumorigenic cells results in an									
XX	CC	inhibition of the tumorigenic properties of the overproducing									
XX	CC	cells. Antagonists to GP88 are used to treat diseases associated									
XX	CC	with increased expression of GP88, particularly cancer but also									
XX	CC	viral infections. Fragments of GP88 are used to raise specific									
XX	CC	antibodies (used as antagonists, as diagnostic reagents and for									
XX	CC	delivering toxins or other compounds to GP88-expressing cells) and									
XX	CC	to screen for antibodies. Methods are provided for diagnosing									
XX	CC	disease, or determining susceptibility to disease, resulting from									
XX	CC	altered GP88 activity.									
XX	XX	Sequence	593 AA;								
XX	XX	Query Match	92.3%;	Score	325;	DB	20;	Length	593;		
XX	XX	Best Local Similarity	91.2%;	Pred.	No. 1.9e-22;						
XX	XX	Matches	52;	Conservative	2;	Mismatches	3;	Indels	0;	Gaps	
XX	XX	AC	AAE20521;								
XX	XX	DT	01-JUL-2002	(first entry)							
XX	XX	DE	Human granulin/epithelin precursor (GP88) protein.								
XX	XX	KW	Human; granulin precursor: GP88; cytostatic; tumourigenicity; tamoxifen								
XX	XX	KW	antineoplastic; antioestrogen therapy; skin cancer.								
XX	XX	OS	Homo sapiens.								
XX	XX	Key	Location/Qualifiers								
XX	XX	FT	Misc-difference 273								
XX	XX	FT	/note= "Encoded by AAG"								
XX	XX	FT	346...364								
XX	XX	FT	Region								
XX	XX	FT	/note= "Region used as antigen to develop anti-human								
XX	XX	FT	GP88 neutralising antibody"								
XX	XX	PN	US2002025543-A1.								
XX	XX	PD	28-FEB-2002.								
XX	XX	PF	15-JUN-2001; 2001US-0880842.								
XX	XX	PR	23-MAY-1997; 97US-0863079.								
XX	XX	PR	08-DEC-1999; 99US-0456886.								
XX	XX	PA	(SERK/) SERRERO G.								
XX	XX	PI	Serrero G;								
XX	XX	DR	WPI; 2002-267529/31.								
XX	XX	DR	N-PSDB; AAD32850.								
XX	XX	PT	Diagnosing tumorigenicity in a human, comprising obtaining a cell								
XX	XX	PT	sample, detecting GP88 in the cells, and determining the number of								
XX	XX	PT	positive cells in the sample -								
XX	XX	PS	Disclosure; Fig 9B; 50pp; English.								

CC The invention relates to a method for diagnosing tumorigenicity in a
CC human. The method comprises obtaining a biological sample containing
CC cells from the patient, detecting GP88 in the cells of the sample,
CC and determining the number of GP88 positive cells in the sample, and
CC determining the ratio of GP88 positive cells to the total number of cells
CC in the sample. The invention also relates to a method for determining
CC if a human patient is resistant to the antineoplastic effects of
CC antioestrogen therapy. The method is useful for diagnosing
CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple
CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,
CC colon, or skin cancer. The method can be used to treat or prevent
CC re-occurrence of cancer in a patient, by administering tamoxifen if the
CC sample contains less than 10 % GP88, or less than 5 % GP88 positive
CC cells. The present sequence is human granulin/epithelin precursor (GP88)
CC protein.

XX
SQ Sequence 593 AA;

Query Match 92.3%; Score 325; DB 23; Length 593;
Best Local Similarity 91.2%; Pred. No. 1.9e-22;
Matches 52; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 IGCDOHTSCPVGTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARTCEKQ 57
DB 442 IGCDOHTSCPVGTCCPSLKGSWACCOLPHAVCCEDRQHCCPAGYTCNVKARSCEKE 498

Search completed: April 21, 2003, 14:20:39
Job time : 36 secs

THIS PAGE BLANK (USPTO)